A Novel Approach for Information Hiding in Punjabi Language

Dissertation submitted to the Central University of Punjab

For the award of

Master of Technology

In

Computer Science and Technology

BY

ARUN KUMAR

Supervisor

Dr. Amandeep Kaur

Centre for Computer Science & Technology
School of Engineering & Technology
Central University of Punjab, Bathinda
September 2016

DECLARATION

I declare that the dissertation entitled " A Novel Approach for Information Hiding in Punjabi Language" has been prepared by me under the guidance of Dr. Amandeep Kaur, Associate Professor, Centre for Computer Science and Technology, School of Engineering and Technology, Central University of Punjab.

No part of this dissertation has formed the basis for the award of any degree or fellowship previously.

ARUN KUMAR

Centre for Computer Science and Technology,

School of Engineering and Technology,

Central University of Punjab, Bathinda - 151001.

Date:

CERTIFICATE

I certify that ARUN KUMAR has prepared his dissertation entitled “A Novel Approach for Information Hiding in Punjabi Language”, for the award of M.Tech degree of the Central University of Punjab, under my guidance. He has carried out this work at the Centre for Computer Science and Technology, School of Engineering and Technology, Central University of Punjab.

Dr. Amandeep Kaur

Associate Professor

Centre for Computer Science and Technology,

School of Engineering and Technology,

Central University of Punjab, Bathinda - 151001.

Date:

ABSTRACT

A Novel Approach for Information Hiding in Punjabi Language.

Name:	Arun Kumar
Registration Number:	CUPB/M.Tech-CS/SET/CST/2014-15/05
Degree for which submitted:	Master of Technology
Name of Supervisor:	Dr. Amandeep Kaur
Center:	Computer Science & Technology
School of Studies	Engineering & Technology

Keywords: Jaro-Winkler, LRM, RLM, NetBeans, Embedding Process, Extraction process, Stego-Key, Space Ratio.

The Internet is global in scope and rapidly growing. With this growth, internet security threats are also increasing. The Internet is the global information exchange media, which is open and insecure, needs more and better security provision. In the era of computer network virtually all business, government and academic organizations interconnect their local area network with a collection of interconnected networks. Data transmission in a public communication system is not secure because of interception and improper manipulation by an opponent. Therefore, the attractive solution for this problem is steganography, which is the art and science of writing concealed messages in such a way that no one, apart from the sender and intend recipient, can notice the existence of the message. Thus, confidentiality and integrity principle of security can be achieved. There have been wide ranges of algorithms introduced using a text file as cover media. This thesis explains a two layered hiding technique that provides high security than other techniques. The use of local language ensures more security to the information exchange as the awareness of the local language is limited. The proposed approach uses the local language Punjabi as cover text. In proposed approach, the cover media is pre-existing media, not system generated media. Therefore, a user will be free to use the cover media taken from any Punjabi newspaper, magazine or book and the syntax of a sentence and sequence of sentence both will be true grammatically. Even though the opponent has a very good command on the Punjabi language, they will not found any suspicious.

Arun Kumar

Dr. Amandeep kaur

(Supervisor)

ACKNOWLEDGEMENT

It is a great pleasure that, I record my indebtedness to my supervisor **Dr. Amandeep Kaur** (Coordinate of Center for Computer Science and Technology), Central University of Punjab, for her excellent guidance and support in completing the dissertation report.

I would also like to extend my sincere and extreme gratefulness to **Prof. A.K. Jain** (Dean, School of Engineering & Technology) Central University of Punjab for his counsel and guidance.

A special thanks to my colleagues in the program, for their support and for making this study period easy and enjoyable.

Last but not the least this acknowledgment will be incomplete if I did not mention the name of my wife, **Sarita**, for motivation, for assuming more than a fair share of household burdens and for her support and patience!

ARUN KUMAR
(CUPB/M.Tech-CS/SET/CST/2014-15/05)
Centre for Computer Science and Engineering
Central University of Punjab

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Sr. No	Full Form	Abbreviation
1	American Standard Code for Information Interchange	ASCII
2	Cascading Style Sheets	CSS
3	Data Encryption Standard	DES
4	Hyper Text Markup Language	HTML
5	Integrated Development Environment	IDE
6	JavaServer Pages	JSP
7	Java Virtual Machine	JVM
8	Longest Common Subsequence	LCS
9	Left-to- Right Remark	LRM
10	Microsoft Compound Document File Format	MCDFF
11	Optical Character Recognition	OCR
12	Probabilistic Context-Free Grammar	PCFG
13	Right-to-left Remark	RLM
14	Structured Query Language	SQL

CHAPTER 1
INTRODUCTION

In this era the internet has global scope, but this global inter-network is open and insecure. The internet the global information exchange media needs more and better security provision. There are numbers of security mechanism like cryptography, digital signature, steganography, etc. “Steganography” is the security mechanism which hides the existence of the communication. In contrast to cryptography, where the attacker can detect, intercept and modify messages without violating certain security property assured by a cryptosystem. But the goal of steganography is to hide messages inside other harmless messages in such a way that no attacker is able to detect that there is a second message present. So steganography technologies are an important part of the future of internet security and privacy. The graphical representation of the steganography process is presented as follows.

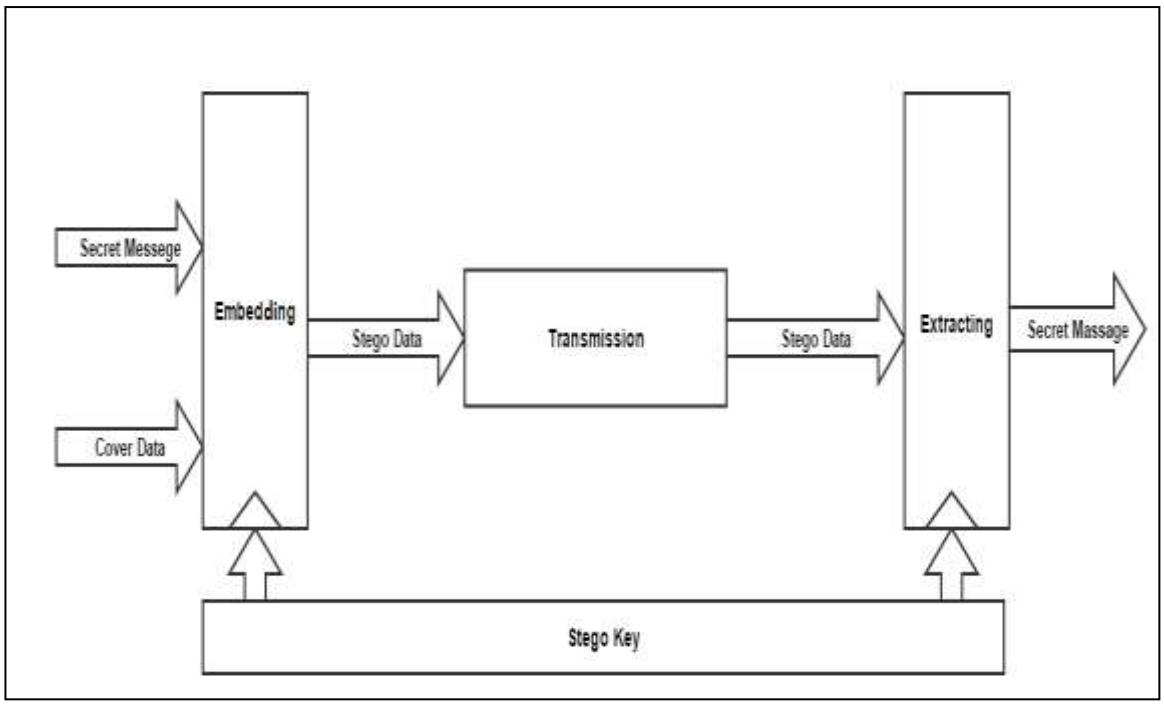

Figure 1.1: Model of Steganography

Secret communication scheme was known a long time ago; Julius Caesar used cryptography to encode his political directions (Singh, Singh, & Saroha, 2009). Steganography is the process of hiding data inside other data in such a way that no one apart from the intended recipient and sender knows the existence of the message. The word steganography is derived from the Greek words “*stegos*”

meaning “cover” and “*grafia*” meaning “writing” (Morkel, Olivier, & Africa, 2005) defining it as “covered writing”. The practice of sending secret messages is not new; it has been made for millennia. Modern steganography is generally referred to deal with electronic media. Steganography consists of two algorithms, one for embedding and one for extracting. For hiding a secret message within a cover at the sender side, an embedding process is used. The secret message is discovered at the receiving end through the extracting process**.**

1.1 Steganography Terminologies

The terminology used while discussing steganography / steganographic Systems is as follows:

- **Message:** Secret Information to be hidden inside the container.
- **Cover Text:** The carrier / media used for hiding a message inside its body.
- **Embedding:** The process of hiding secret information inside Cover.
- **Stego Object:** The resultant Cover after going through embedding.
- **Extraction:** The process of extracting hidden information from Stego Object.
- **Stego Key:** Secret Key used in information embedding and extraction.
- **Stego System:** Figure 1.1 depicts entire activity of embedding secret information inside some Cover by employing a Stego key at the sender’s end and its extraction using the Stego key at the receiving end, and constitutes a complete Steganographic System.
- **Redundant Bits:** Pieces of information inside a file which can be overwritten or altered without damaging the file.

1.2 Characteristics of Steganography

Though the most obvious goal of steganography is to hide data, but there are several characteristics of steganography used to judge the strength and weaknesses of steganography methods. Mainly there are three characteristics of steganography (Codr, 2009) as mentioned below:

- **Capacity:** This feature specifies how much data can be hidden in such a way that, it should be possible for humans to detect a distortion in the stego-object.
- **Un-detectability**: This feature specifies inability for a computer to use statistics or other computational methods to differentiate between covers and stego-objects.

- **Robustness:** Robustness of the steganography refers to the ability of the message to persist regardless of compression or other common modifications.

These three main components work in opposition to one another. If one of these is increased, it causes the others to decrease. Hence, no steganographic technique can be perfectly undetectable and robust and have the maximum capacity (Salommon, 2003). In most cases, robustness and undetectability are most important as compared to capacity whereas watermarking favors robustness most strongly. In general, the most of the steganography methods considers undetectability most important. The balance of the properties of good steganography is presented in the Fig. 1.2.

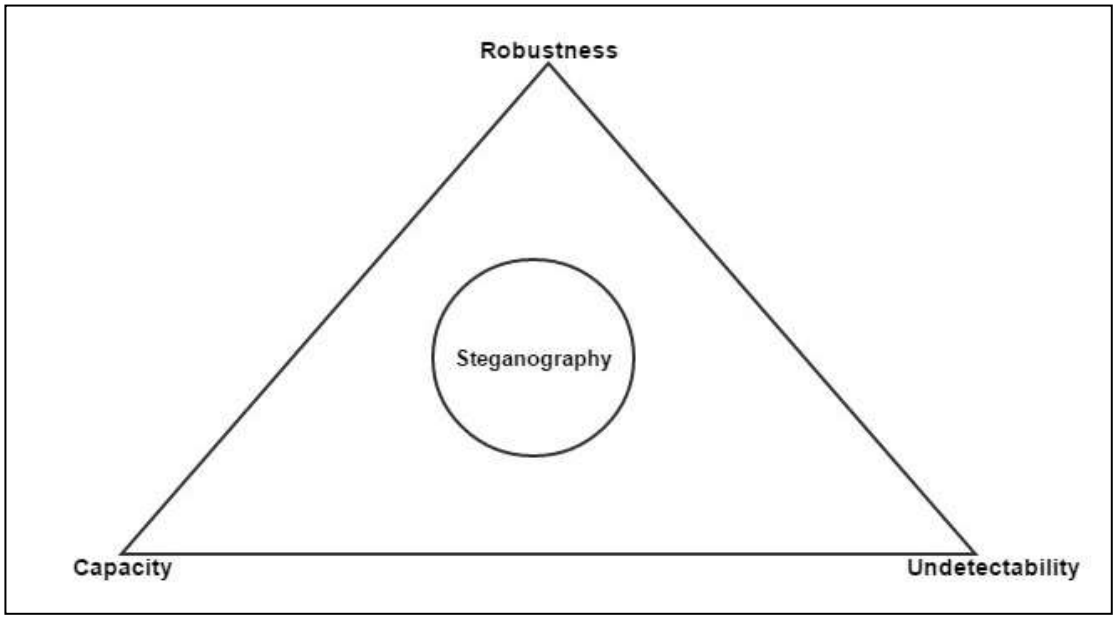

Figure 1.2: Properties of Steganography

1.3 Types of Steganography

It is possible that the graphic or sound files can be altered slightly without losing their overall viability for the viewer and listener. With audio, you can use bits of the file that contain sound not audible to the human ear. With graphic images, redundant bits of color from the image can be removed and still produce a picture that looks undamaged to the human eye and is difficult to distinguish from its original (Abduallah, Rahma, & Pathan, 2014) (Nosrati, 2012). Due to the absence of redundant information in a text file as compared to a picture or sound file, the text steganography is the most difficult type of steganography (Kabetta, Dwiandiyanta, Dwiandiyanta, & Suyoto, 2011). According to categories of file formats that can be

used for cover media, there are four types of steganography techniques:- image, audio, video and text (Bhattach-aryya, Banerjee, & Sanyal, 2011).

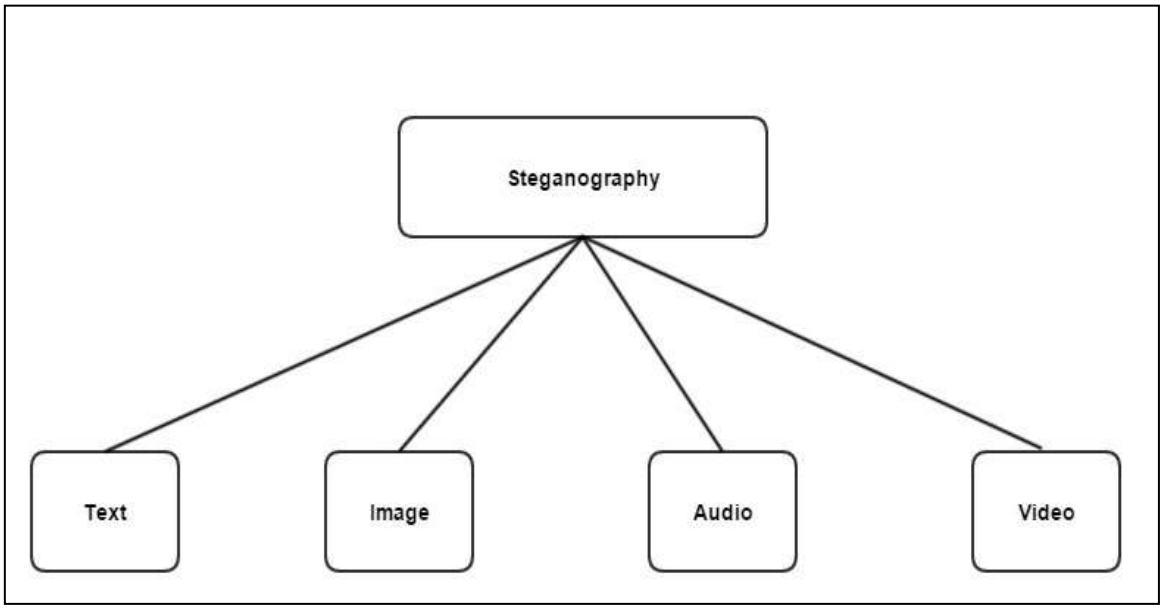

Figure 1.3: Types of Steganography

1.4 Applications of Steganography

Steganography can be used for a range of purposes. Legitimate purposes can include things like covert channels, embedded data and digital watermarking.

I. If Steganography is used to hide data in IP header or TCP header, it is called Covert channels in TCP/IP. In this approach, some fields of IP header or TCP header are selected for data hiding. In this way, absolute secrecy is maintained on the internet for an entire communication process and not just one document.

II. Today the most of the popular use of steganography is to hide the secret message by embedding it into cover media. This method of steganography is very useful when a party must send top secret, private or highly sensitive documents over an open systems environment such as the internet. It is possible to send news and information without the fear that the messages being intercepted and traced back (Kumar & Pooja, 2010) (Krenn, 2004).

III. Steganography can also be used to implement watermarking. Although the concept of watermarking not a pure steganographic technique, but several steganographic techniques that are being used to store watermarks in data. Company or entities usually use digital watermarking for Copywrite reasons to safeguard their property by either embedding their license information into their property or by hiding trademark. Digital Watermarking is very important in the

detection and prosecution of software pirates and digital thieves (Kumar & Pooja, 2010) (Krenn, 2004).

IV. E-commerce permits an interesting use of steganography. In current e-commerce transactions, most of the users are secured by a username and password.There is no real method of verifying that whether the user is the actual cardholder or not. By embedding the unique session IDs into the biometric fingerprint images via steganography, allow for a very secure option to open E-commerce transaction verification (Krenn, 2004).

Unfortunately, Steganography can also be used for illegal motives. For instance, if someone is trying to steal data, they could hide it in another file or files and mail it out in an innocent looking email or file transfer. Furthermore, a person with a hobby of saving images and video clips to their hard drive may choose to hide the proof through the use of steganography as it was concerned for terroristic purposes. Hence, steganography can be both a legitimate and illegitimate application (Abbadi, 2008).

1.5 Text Steganography

Text steganography uses text as the cover medium in which information is hidden. Hiding information in plain text can be done in many different ways.

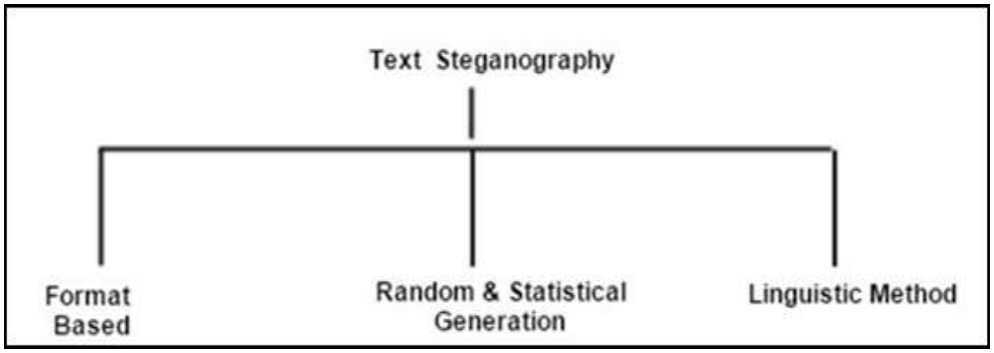

Figure 1.4: Type of Text Steganography

Text steganography can be basically classified into three types- format-based, random and statistical generation and Linguistic method (Bhattacharyya *et al*., 2011). Within each category, the text can be either system generated or embedded within known plaintext.

1.5.1 Format-Based Methods

This type of method uses and changes the formatting of the cover-text to hide data. They do not change any word or sentence, so it does not harm the value of the

cover-text. Format-based methods use physical text formatting of text as a place in which to hide the information. Generally, in order to hide the steganographic text, this method modifies existing text. Deliberate misspellings distributed throughout the text, insertion of spaces, resizing the fonts are some of the many format-based methods being used in text steganography. A format-based text steganography method is open space method (Bhattacharyya *et al.*, 2011). In this method, extra white spaces are added into the text to hide information. Another two format-based methods are word shifting and line shifting. In word shifting method, the horizontal alignments of some words are shifted by changing distances between words to embed information. In line shifting method, vertical alignments of some lines of the text are shifted to create a unique hidden shape to embed a message in it.

1.5.2 Random and Statistical Generation

It is generating cover text according to the statistical properties. This method is based on character sequences and words sequences. Information hiding in Random Character and word sequences method (Bennett, 2004) generates a random sequence of characters or words to hide the information.

A second approach is to generate cover-text automatically according to the statistical properties of language. These methods use predefined grammars to produce cover-text in a certain natural language. A probabilistic context-free grammar (PCFG) is a commonly used language model where each transformation rule of a context- free grammar has a probability associated with it (Bhattacharyya *et al.,* 2011). The sentences of cover text are constructed according to the secret message to be hidden in it. The quality of the generated stego-message depends directly on the quality of the grammars used.

1.5.3 Linguistic Method

The linguistic method (Salommon, 2003) hides the secret message by modifying the text accordance to linguistic properties of the text. In this method, the linguistic structure of the text is used as a place to hide information. The syntactic method is a linguistic steganography method where some punctuation signs like a comma (,) and full-stop (.) are placed in proper places in the document to embed a data. This method needs proper identification of places where the signs can be inserted. This approach is vulnerable to attack because the inconsistent use of punctuation is noticeable.

Another linguistic steganography method is a semantic method. In this method, some pre-selected synonyms of the word are used. The words are replaced by their synonyms to hide information in it. By using a special word, data are embedded in the text. The sender and receiver using such a method may agree on the use of a certain online thesaurus.

1.6 Punjabi Language and its Orthography

The Punjabi language is one of the members of the Modern Indo-Aryan family of languages, also known as Indic languages.There are 122 languages in India; among these 22 languages have been declared as official languages by Government of India . According to 2011 Census of India, there are 27,704,236 Punjabi speakers in India. The Punjabi Language is spoken and used in both parts of Punjab, in India and also in Pakistan. Punjabi in India is heavily influenced by Sanskrit, while in Pakistan, Punjabi is heavily influenced by Perso-Arabic. Punjabi uses two different scripts.

I. **Shahmukhi** is a Perso-Arabic script used by Muslims of Pakistan and it reads from right to left. “Shahmukhi” means “from the mouth of the kings.”(Veach & Williamson, 2015)
II. **Gurmukhi** is the script used by Hindus and the Sikhs of Indian. “Gurmukhi” means “from the mouth of the Gurus”. Gurmukhi is the script used by the Sikh Gurus and descended from the Brahmi script. Gurmukhi script is used primarily for the Punjabi language. It is written and read from left to right (Veach & Williamson, 2015) (Lehal & Bhagat, 2004).

1.6.1 Unicode of Gurmukhi

The Unicode Standard is a character coding system designed to support the worldwide interchange, processing, and display of the written texts of the diverse languages and technical disciplines of the modern world. In addition, it supports classical and historical texts of many written languages.

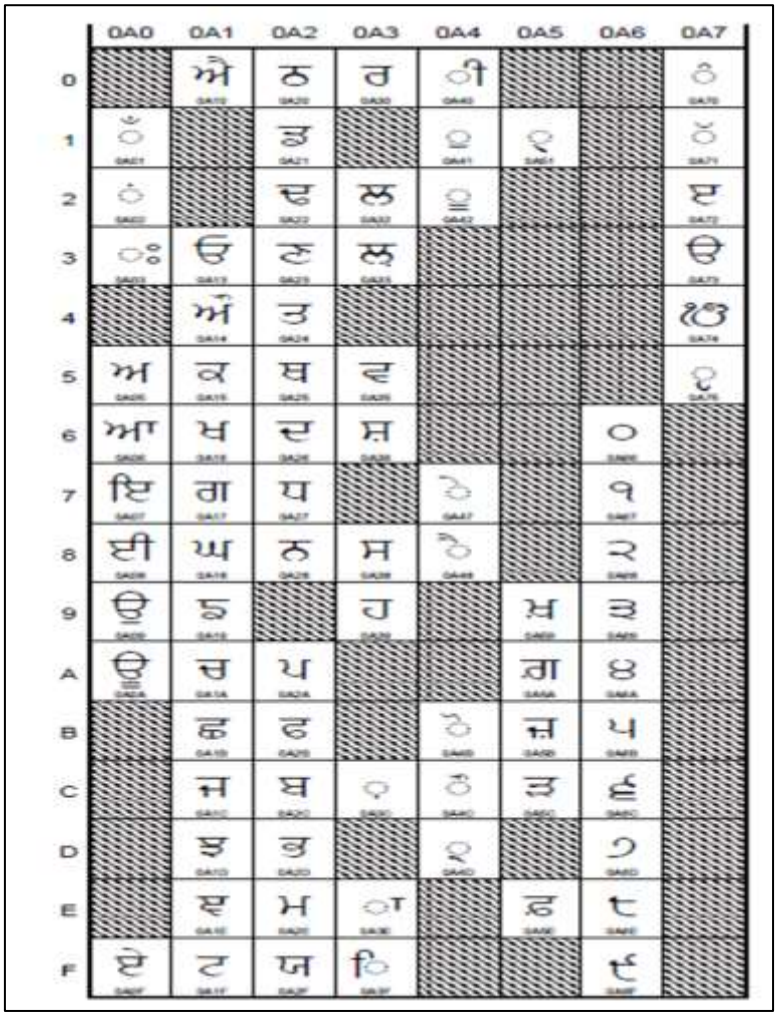

Figure 1.5: Unicode of Gurmukhi Characters (Unicode consortium, 2015)

Table 1.1: Consonants of Punjabi

ਕ	ਠ	ਬ
ਖ	ਡ	ਭ
ਗ	ਢ	ਮ
ਘ	ਣ	ਯ
ਙ	ਤ	ਰ
ਚ	ਥ	ਲ
ਛ	ਦ	ਲ਼
ਜ	ਧ	ਵ
ਝ	ਨ	ਸ਼
ਞ	ਪ	ਸ
ਟ	ਫ	ਹ

Table 1.2: Additional Consonants of Punjabi

ਖ਼	ਜ਼	ਫ਼
ਗ਼	ੜ	

Table 1.3: Independent Vowels of Punjabi

ਅ	ਉ	ਓ
ਆ	ਊ	ਅੌ
ਇ	ਏ	
ਈ	ਐ	

Table1.4: Dependent Vowels of Punjabi

ਾ	ੁ	ੈ
ਿ	ੂ	ੋ
ੀ	ੇ	ੌ

Table 1.5: Additional Characters of Punjabi

ੱ	੍	ੲ
ੰ	਼	ੳ
੃	ੁ	ੴ
੍	ੱ	۔

CHAPTER 2

LITERATURE REVIEW

Text steganography is an important area under information hiding. This thesis work focuses on developing information hiding in text using the Punjabi language. A complete review of the text steganography is as given below:

Most research of steganography has been performed using cover media such as images, video clips and sound. In addition to other types cover media, text can also be used as cover media. The advantages to prefer the text as a cover media over other media is its lesser memory occupation and needs low bandwidth (Shirali-Shahrez & Shirali-Shahreza, 2006). Some of the methods proposed to hide the information in the cover text are line shifting, words shifting, whitespaces manipulation, feature coding, syntactic and semantic method.

(Low *et al.* 1995) proposed a method to hide a watermark by vertically shifting the position of the locations of text lines i.e. Certain text-lines are shifted slightly up or down from their normal positions. In word-shift, a block of words is shifted slightly to the left or right of their normal position to hide the mark. This technique can also be used to hide steganographic information. If the text is retyped or if character recognition programs (OCR) are used, the hidden information would get destroyed. (Bender *et al.,* 1996) proposed a method in which whitespace between words are manipulated to encode 0 and 1. One space between words is interpreted as a “0.” Two spaces are interpreted as a “1. The drawback of this method is that it requires larger spaces to encode a single character and data hidden is destroyed once the spaces are deleted. Some of the problems appeared in Bender’s method has been solved by (Por & Delina, 2008). (Por & Delina) proposed a method in hiding information using inter-words spacing and line spacing between paragraphs as a hybrid method. In this way, most of the whitespaces in text document have been utilized. The unique feature of the method is to generate a cover-text dynamically according to the length of the secret message. But still there is a drawback that the hidden data is destroyed once the spaces are deleted, by some word processing software.

Some researchers conducted research on feature coding method. (Shirali-Shahreza & Shirali-Shahreza, 2006) introduced a method for text steganography using feature

coding method. In this approach, Persian/Arabic language has been used. In Persian 18 letters out of 32 alphabet letters have points and in Arabic, 15 letters out of the entire 28 alphabet letters have a point. Data is hidden by considering the existence of too many points in Persian and Arabic text. The information is hidden in the text by vertical displacement of the points. If Characters with point remains unchanged, then it represents 0. The bit 1 is represented by shifting the point on the concerned character a little upward. (Changder *et al.,* 2009), presented a new approach for steganography in Hindi texts by using feature coding method. In this approach, the secret message is hidden in the text by shifting the specific matra towards left or right. The advantage of these methods is that a large volume of information can be hidden in the text. But the information hidden is lost in case of retyping.

Some scholars presented research on a semantic method of text steganography. (Shirali-Shahreza, 2008) proposed a method based on the semantic approach of text steganography to exploit same words which are spelled differently in British and American English for hiding secret message bits. In English some words spelled in British and American English is different. This differences in spellings form the basis of steganography. This method has little capacity to hide data in the text. However, it is dependent on the body of text and its size, but overall its capacity is very low.

Table 2.1. Different Spelling in American & British Language

American Spelling	British Spelling
Favorite	Favourite
Criticize	Criticise
Fulfill	Fulfil
Center	Centre

(Shirali-Shahreza & Shirali-Shahreza, 2008) proposed a new method based on the semantic approach in which hiding capacity has been improved. This method uses the synonyms of certain words thereby hiding information in the text. The synonyms substitution may represent a single or multiple bit combinations for the secret information. A major advantage of these methods is the protection of information in case of retyping or using OCR programs. However, the later method may alter the meaning of the text.

Table 2.2.Some Synonym Words

Big	Large
Small	Little
Chilly	Cool
Smart	Clever
Spaced	Stretched

Limited works have been done on hiding information in Hindi text. (Alla & Prasad, 2008) proposed a method for the Hindi language. The Hindi language consists of letters and letter diacritics. In this method, the bit 0 is encoded with vowel and consonant, and in the same way, the bit 1 is hidden by letter diacritics. The stego-text is system generated and it may be irrelevant text to be transferred. This irrelevant text is an advantage as well disadvantage because an attacker must have a full command over the language and its structure to decode. But if an attacker has the command over language this irrelevant text has been suspected containing a secret message.

(Changder *et al*, 2010) proposed a method to hide the secret message in the text through Indian languages. The Indian languages have much more characters as compared to other languages. Therefore, it will be quite easy to create the words or innocent sentences using Indian languages. In this approach cover text is system generated. The Information hiding is done by creating meaningful sentences after finding the longest common subsequence of two binary strings among which, one is the secret message and another may be any binary string. This study provides large hiding capacity text steganography with minimal change of the structure of the cover file. Therefore, it will be less noticeable to the unintended recipients. But through the LCS of two binary string, it is not so easy to create the meaningful sentences always.

There are some other methods for text steganography which uses internet programming language file as cover text.(Garg, 2011) presented a text Steganography by using HTML files. Author classified HTML into two categories: primary attributes and secondary attributes. If a secondary attribute is followed by a primary attribute, then a 0 bit is detected, else a 1 is detected. The author suggested applying two steps. The first step is encryption to improve the message security.

The second step suggests applying HTML Steganography scenarios to hide the bits. In this approach, the HTML file is scanned to search all tags in the web page to classify them into primary attributes and secondary attributes. If the hidden bit is 1, the primary and corresponding secondary attribute are swapped. Otherwise, no change is applied. The main advantage of this algorithm is that all the changes in the code file result in no effect on the web information. After the development of XML, most of the web pages code comprises of XML code. (Inoue *et al.,* 2001) represented the mechanism for hiding the message in the XML file. In this approach, an empty element is used to hide the 0 and 1. The empty element can be either a start-tag immediately followed by an end-tag or an empty-element tag. By using these two form of empty element bit 0 and 1 is concealed as follows:

stego key:

<img></img> … 0

<img/> … 1

stego data:

<img src=”foo1.jpg”></img>

<img src=”foo5.jpg”></img>

Embedded data:

0110

Cascading Style Sheets (CSS) is a stylesheet language which is used to describe the presentation semantics of a document written in HTML and XHTML. (Kabetta & Dwiandiyanta, 2011) proposed a new scheme for hiding information in CSS file. In this scheme, a CSS file is used as a cover media to embed messages. The secret message to be hidden in CSS files is first encrypted using RSA as a public key cryptographic algorithm and it is applied to the CSS file. The binary bits are represented by white spaces either spaces or tabs after semi-colon characters of each CSS instruction. The approach creates the stego-text that looks same as the original text, by using the “End of Line” techniques for the embedding process and there are no obvious changes. The disadvantage of this technique is the limited

amount of characters that can be embedded. The embedding capacity depends on the available amount of semicolons in the text.

Many types of steganography were in use; however, there was no known steganography method for query languages such as SQL. (Bassil, 2012) proposed a new scheme for hiding the information using SQL Queries. This scheme is system generated method in which, SQL queried is generated according to secret message to be hidden. There is a dictionary of words organized into 65 categories represent 65 different characters including the 26 letters of the English language, the 10 digits of the decimal system, and a set of 29 special characters with no common words between these categories. For each character in the secret message, a word in matched category is selected randomly. All generated words are then split into two parts; first part represent the terms of SELECT clause and second part represents the terms of WHERE clause. In this way cover media which is SQL query is generated. The advantage of this scheme is that different output SQL carrier will be generated for the same input secret message by changing the content of the dictionary from time to time and it is multilingual. The disadvantage of this scheme is that the cover media is system generated due to which it is possible that the generated query may not be meaningful SQL query.

Besides all these the other steganography approaches have also been developed. (Rahma *et al.,* 2013) presented new scheme for hiding the information in the English language based on Unicode. Many English alphabet characters have been used in another language, with different codes and different glyph. A secret message is concealed by replacing the character in a cover file with the same character in another language having different Unicode and different glyph.

Table 2.3. Unicode of Characters in Multilingual

English Alphabets and ASCII code	**Multilingual, Glyphs, and Unicode**		
A 0041	Greek	Cyrillic	Cherokee
	Α	А	Α
	0391	0410	13AA

The advantage of this scheme is the cover media is pre-existing. The disadvantage of this scheme is that the size of the steo-Text file is increased and hiding capacity depends on the frequency of selected characters used in the cover text.

(Agarwal, 2013) proposed a new scheme for hiding the information using start and end letter of the words of a cover file. Bit 0 or 1 is hidden by reading a word sequentially from the cover file and including the starting letter or the end letter, respectively, of the word in the stego key. In this way, the secret message is concealed without any change made to cover media. The advantage of this scheme is highly robust because no font-altering or feature encoding is used. The disadvantage of this scheme is every time both the stego- text and stego-Key have to be sent to the receiver.

(Uddin *et al*., 2014) proposed a new scheme for hiding the information along with DES cryptography. The cover text has been prepared as ordinary as possible. The secret message to be hidden on cover files is first encrypted using DES as a symmetric key cryptographic algorithm and it is applied to the cover file. The each encrypted character is counted for the frequency with respective position and finally, it compared with a cover text to retrieve the position in the cover text. Now the position of encrypted in cover media is converted to equivalent ASCII code. In this way, the message is hidden in the cover text. The advantages of this scheme are its high security and cover media is not system generated. The disadvantage of this scheme is it is detectable due to adding puzzle at the end of cover media.

(Dasare & Dhore, 2015) proposed a new scheme for hiding the information in MS Word Document Using MCDFF. MCDFF is a compound file of MS word document. A compound file provides a solution to store efficiently multiple kinds of objects in one document. It is nothing but file system within a file. It is used to find out the number of unused blocks available in the document file. This unused block is used to hide the secret message. The advantage of this scheme is that the secret message is hidden in such a way that it is independent of the contents of the document file. The disadvantage of this scheme is that more raw blocks have to be added in the cover document if a secret data size is too large.

2.1 Summary of Literature Review

A variety of methods for hiding the secret message in the text has been developed

so far with different languages. No single steganography technique is there which can be used for all languages, because of their different structure and specification. Some of them have either lack in robustness or stego-text is system generated. Still some techniques face less hiding capacity issue while some other techniques are easily noticeable. So there is still a room for designing a robust technique for embedding the secret information into the text media.

Table 2.4. Overview of Literature Review

S. No.	Author /Paper	Parameters/Approach	Advantage	Disadvantage
1	(Shirali-Shahreza, M., 2008	Different spelling in British and American English	High robustness	Less hiding capacity
2	(Shirali-Shahreza, M. H., & Shirali-Shahreza, M., 2008	Substitution of synonym	High robustness	Meaning of text can be altered
3	(Por, L. Y., & Delina, B., 2008)	Whitespaces manipulation between words	High capacity	Hidden data is lost once the spaces are deleted.
4	Alla, K., & Prasad, R., 2008),	Letters and letter diacritics in the Hindi language	Irrelevant Text generated	Irrelevant Text generated
5	Changder, S., Debnath, N. C., & Ghosh, D., 2009	Feature encoding in the Hindi language	High capacity	Not robust

(Cntd..)

Table 2.4. (Cntd..)

6	(Changder, S., Ghosh, D., & Debnath, N. C., 2010)	LCS based system generated cover text in the Hindi language	High capacity	Irrelevant Text generated
7	(Kabetta, H., & Dwiandiyanta, B. Y., 2011),	Places after Semicolon used in CSS file.	Robustness and unnoticeable	Less hiding capacity
8	(Bassil, Y., 2012)	Hiding the information using SQL Queries.	Different output SQL carrier for the same input secret message	cover media is system generated
9	(Rahma, A. M. S., Bhaya, W. S., & Al-Nasrawi, D. A., 2013)	English alphabet characters used in another language, with different codes and different glyph	Size of steo-Text file is increased	the size of the steo-Text file is increased
10	(Agarwal, M., 2013)	Start and end letter of the words of a cover file is used and stego-key generated	Highly robust	Every time both the stego- text and stego-Key have to be sent to the receiver.
11	(Uddin, M. P., Saha, M., Ferdousi, S. J., Ibn Afjal, M., & Marjan, M. A., 2014)	Hiding along with DES and Counting frequency of character with respective position in secret message	Highly secure and cover media not system generated	Easily noticeable

(Cntd..)

Table 2.4. (Cntd..)

12	(Dasare, A. J. *et al.,* 2015)	Unused block in MCDFF word Document.	the secret message is independent of the contents of the document file	More blocks have to be added in the cover document if a secret data size is too large.

A number of studies and researches regarding text steganography method has been plotted as shown in above table. The methods like feature encoding that is not robust against digital copy-paste operation and retyping the same characters or words and steganographic methods that generates the cover media itself is easily noticeable, the proposed method is able to plug the limitations of these two existing techniques. The proposed method also increased the embedding capacity as evident from table.

CHAPTER 3

PROBLEM STATEMENT AND PROPOSED SOLUTION

3.1 Problem Statement

Steganography can be implemented to secure data using different media as cover media. Text steganography is an effective method to secure data with lesser storage requirement and bandwidth for transmission.

Languages and their structure play an important role in developing the steganographic system. On analysis of Punjabi language compared with English language, it was found that unlike the English language it has certain alphabets that are neither vowel nor consonant. This characteristic of Punjabi language can be effectively used for embedding secret data into Punjabi text. Thus, in this thesis work a text steganography based on the Punjabi language will be developed.

3.2 Objectives

- Study and understand existing text steganography
- Study of characteristics of Punjabi text
- Development of Steganographic system using Punjabi text as a cover media
- To evaluate the performance of the proposed scheme based on capacity, robustness, and invisibility.

3.3 Hypothesis

It has been hypothesized that this proposed information hiding scheme will result in a highly secured text steganography method.

3.4 Proposed Scheme

Unlike the English language, the Punjabi language has some alphabets that are neither vowel nor consonant. Languages and their structure play an important role in developing the steganographic system.

A new approach is proposed for text steganography by designing two stego-keys that will be used for two layer hiding. This approach makes the use of pre-existing any meaningful piece of Punjabi text as cover file to hide the secret message.

3.4.1 Data Flow Diagram

Data Flow Diagram of proposed steganography mechanism is illustrated in Fig.3.1.

Firstly, a secret message will be concealed in a cover text by applying an embedding algorithm with stego-key1. The output of first layer hiding will be the input of the second layer. In second layer embedding, stego-key2 and cover-text will be used. Then generated an output file of the second layer will be transmitted using a transmission channel to a receiver. For extracting the secret message which has been sent by the sender, the receiver needs to use an extraction algorithm which is parameterized by a stego-key. At receiving end, the use of stego-keys will be used in reverse order such as stego-key2 will be used at the first layer and stego-key1 at the second layer. A stego-key is used to control the hiding process so as to restrict detection and parties who know it will be able for recovery of the embedded data.

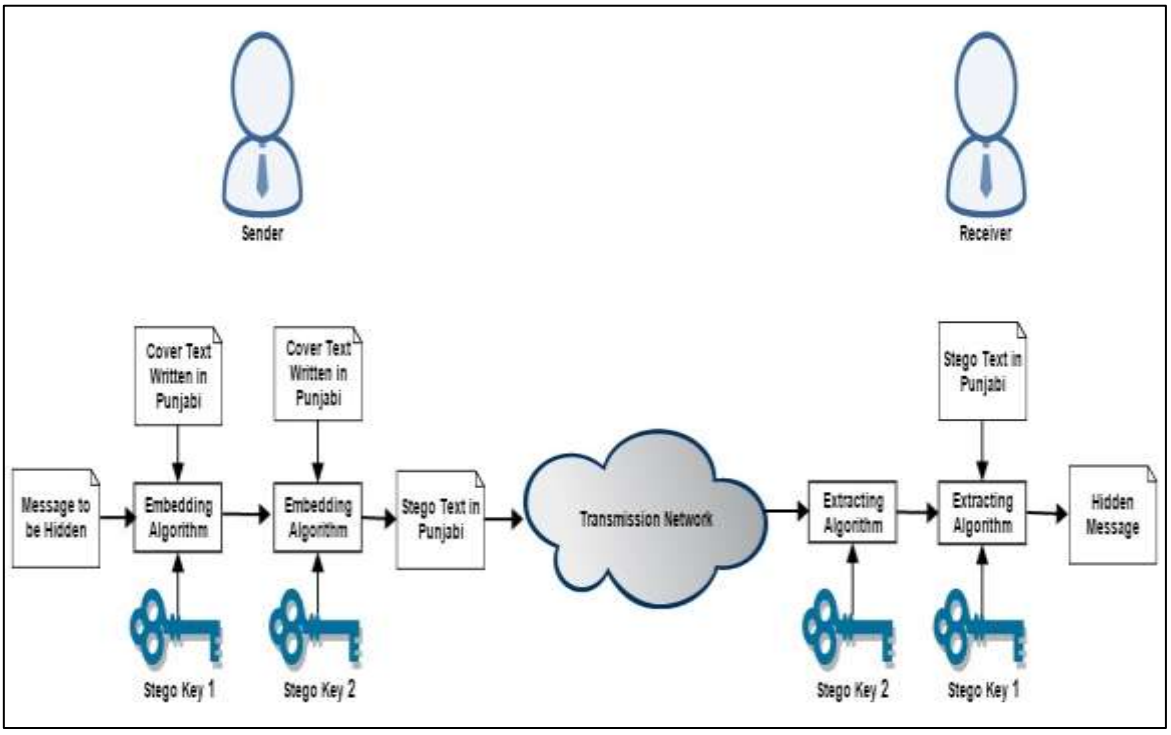

Figure 3.1: Data Flow Diagram

3.4.2 Stego-key1

A property of a sentence can be specified as the number of vowels, the number of consonants, the number of characters and the number of words, now we can assign a value to a sentence as follow:

$$(S) = \begin{cases} 0 & if\ number\ of\ vowels\ is\ even\ otherwise \quad 1 \\ 0 & if\ number\ of\ consonant\ is\ even\ otherwise \quad 1 \\ 0 & if\ number\ of\ characters\ is\ even\ \ otherwise \quad 1 \\ 0 & if\ number\ of\ Words\ is\ even\ otherwise \quad 1 \end{cases}$$

3.4.3 Stego-key2

There are two control characters defined in Unicode Left-to-Right Remark and Right-to-left Remark (Odeh, Elleithy & Faezipour, 2013). These are Nonprinting characters that do not appear as glyphs on the screen. Now Left-to- Right or Right-to-left Remark will be inserted where space character between words as follows:

$$f(x) = \begin{cases} LRM\ \ character\ if\ 0 \\ RLM\ character\ if\ 1 \end{cases}$$

3.4.4 Flowchart of First Layer Hiding at Sender Side

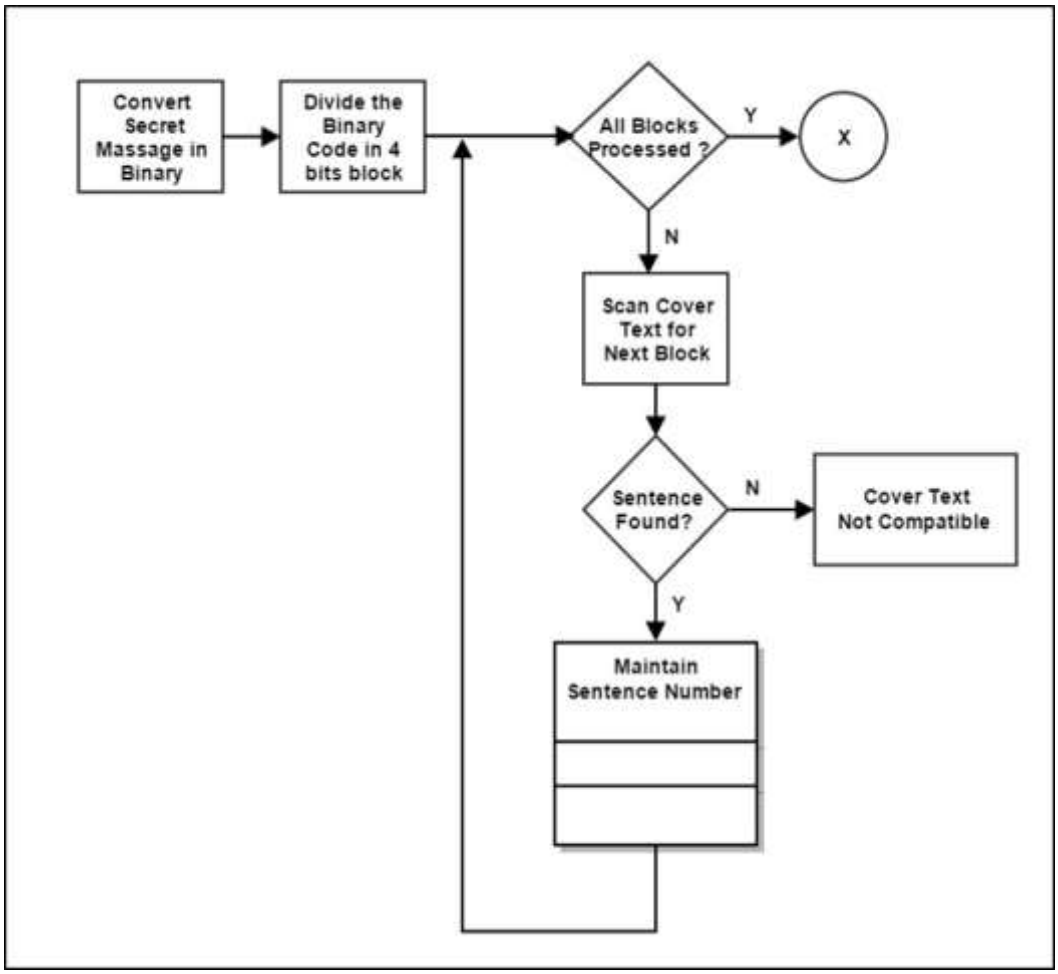

Figure 3.2: First Layer Embedding

To represent a binary stream of 4-bits say '1100' it will select a sentence where the number of vowel, number of consonant, number of character, and number of word are odd, odd, even and even respectively. Therefore to hide a secret message i.e. a stream of binary digits of length N, there should be (N/4) numbers of sentences minimum in cover text. The embedding process is given in Fig. 3.2. Now the

sentence's number of sentences which represent secret message are sent to the second layer of hiding.

3.4.5 Flowchart of Second Layer Hiding at Sender Side

To indicate the sentence number, say 6 and 2, to the receiver in proper sequence it will be concealed in space characters of the cover text by using the stego-key2. The embedding process is given in the Fig. 3.3.

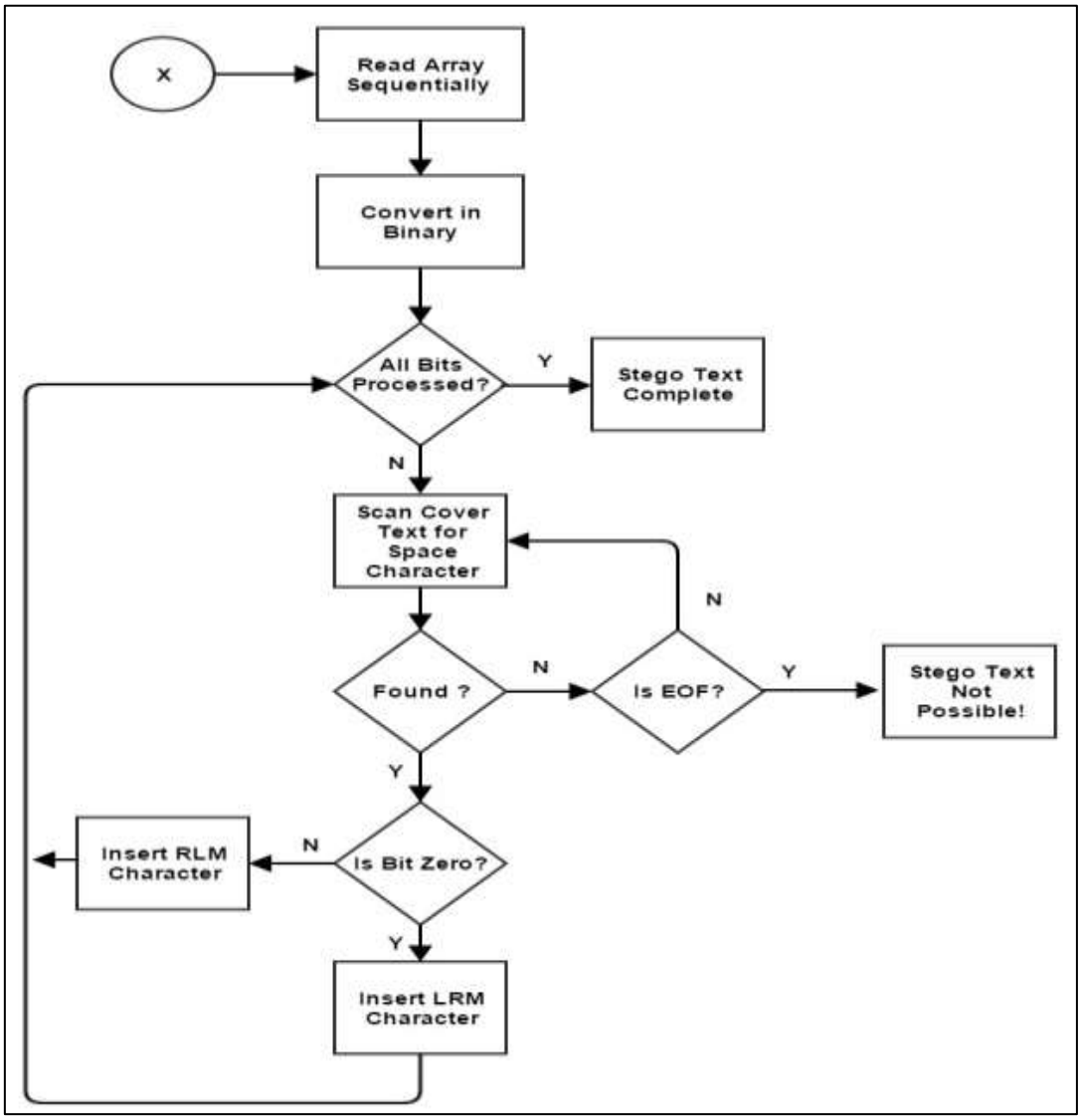

Figure 3.3: Second Layer Embedding

3.4.6 Flowchart of First Layer Extraction at Receiver Side

At the receiver side, to extract the secret message from cover media, first of all sentence's number has to be extracted from the cover text by using stego-key2. The stego-text is scanned for LRM & RLM characters. The whole first layer extraction process is given in Fig. 3.4.

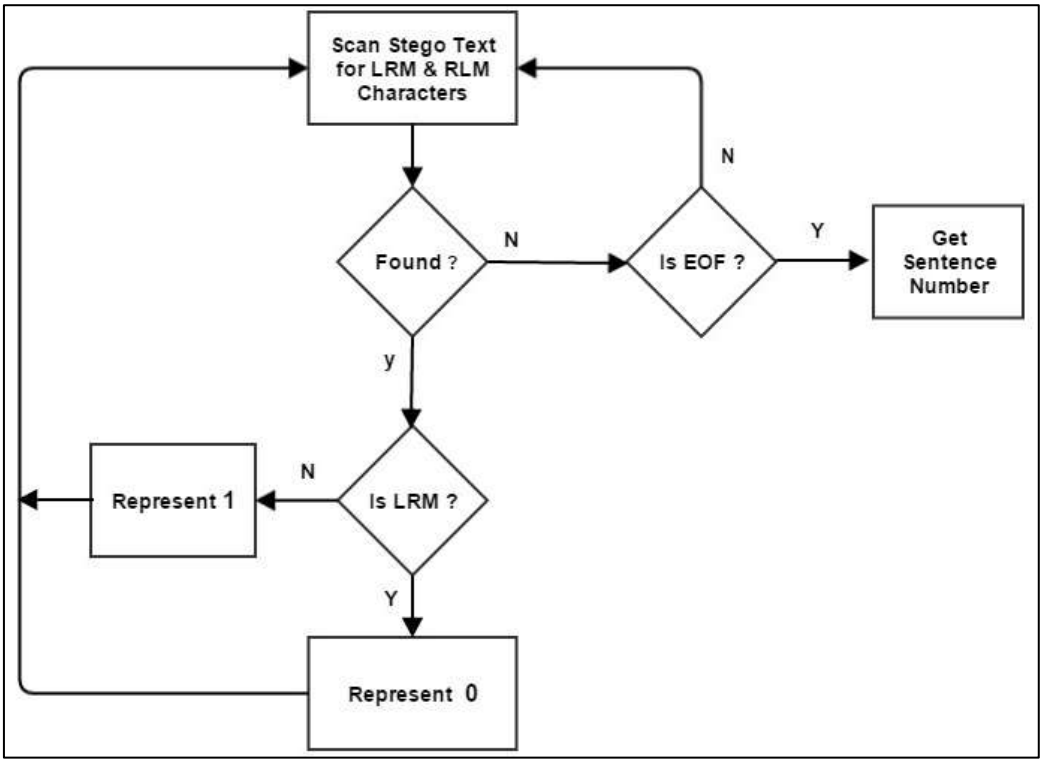

Figure 3.4: First Layer Extraction

3.4.7 Flowchart of Second Layer Extraction at Receiver Side

After extracting the sentence's number, there are two tasks have to be performed, locate the sentences in the cover text containing the block of bits and arrange the block of bits in its proper position, to extract the secret message from cover media. The second layer extraction process is illustrated in Fig. 3.5.

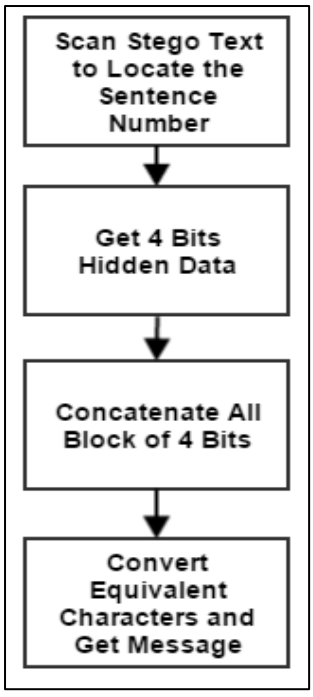

Figure 3.5: Second Layer Extraction

CHAPTER 4

EXPERIMENTAL WORK AND RESULTS

4.1 Software Tools

For simulating the proposed system two software tool (NetBeans Java & Pycharm) have been used. The details of used tools have been given in following sub sections.

4.1.1 NetBeans

NetBeans is an Integrated Development Environment (IDE) for developing primarily with Java, but also with other languages, in particular PHP, C/C++, and HTML5. It is also an application platform framework for Java desktop applications and others. The NetBeans IDE is written in Java and can run on Windows, OS X, Linux, Solaris and other platforms supporting a compatible JVM. The NetBeans Editor editor supports many languages from Java, C/C++, XML and HTML, to PHP, Groovy, Javadoc, JavaScript and JSP.

4.1.2 Pycharm

PyCharm is a graphical Integrated Development Environment (IDE) used for programming in Python. It provides a graphical debugger, code analysis, an integrated unit tester, integration with version control systems and supports web development with Django. PyCharm is developed by JetBrains. It is cross-platform working on Mac OS X , Windows and Linux .The Professional Edition is free for open source projects and for some educational uses.

4.2 Experimental Work

The following sections discuss our experiment results through the example for both Hiding and Extraction of the message. Both hiding and extraction process shows inputs, outputs and the intermediate steps of each process.

4.2.1 Embedding Process

The following Fig. 4.1 shows the graphical user interface at sender side for the embedding process. It accepts the secret message to be concealed, cover text file for hiding secret message inside its body and empty text file where stego-text will be saved. The 1st layer output box shows a list of sentence's number. These sentences have embedded the secret message. After clicking the 2nd layer embedding, the contents of list are embedded in same cover media and stego-text is saved in mentioned file.

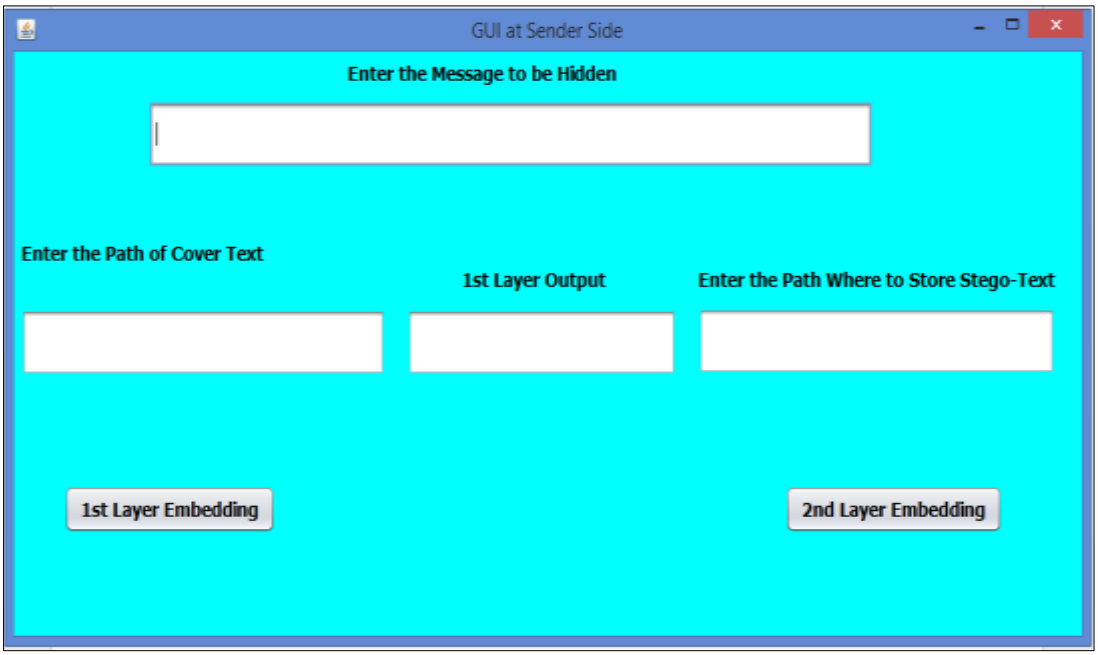

Figure 4.1. GUI for Sender

4.2.1.1 First Layer Embedding

A. Input

Secret Message to be embedded is "**To**".

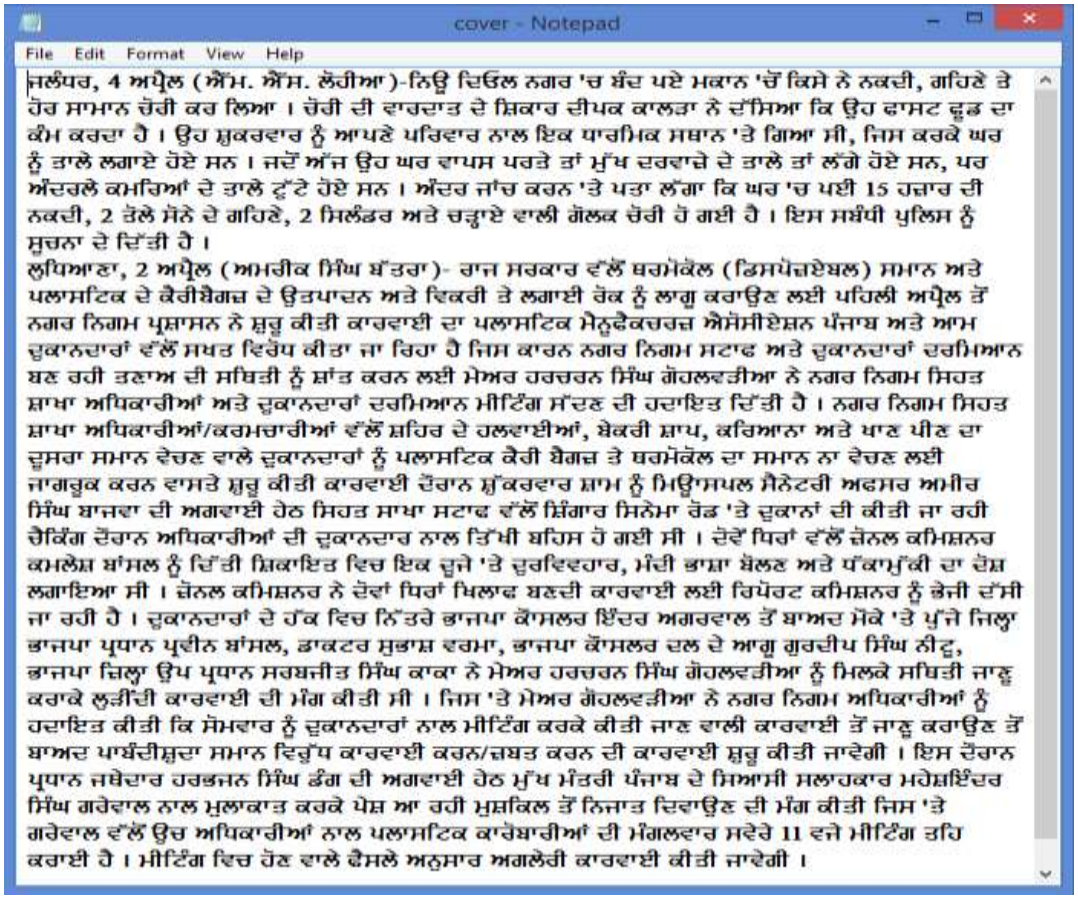

Figure 4.2: Original Cover File

The cover text is a text file named “cover.txt” written in the Punjabi language shown in Fig. 4.2, containing news taken from e-paper “**Ajit-Punjab Di Awaaz**”. If the cover media is not compatible, the message will be displayed to change the cover media.

B. Intermediate Steps

Step 1: Secret message to be embedded is converted into binary form. The binary form of “To” is **0101010001101111**

Step 2: The binary form is divided into 4 bits block as shown below in different colors.

0101010001101111.

Step 3: scan the cover text file “cover.txt” to find the sentence that follows the properties represented by first block 0101. It means that a sentence should contain an even number of vowel, an odd number of consonant, an even number of character and odd number of word. If the sentence is found with desired properties, that sentence’s number in the cover text file is maintained in a list. Similarly, the second, third and fourth blocks are embedded to hide the secret message in step 1. The blocks embedded in sentences and corresponding sentence’s number are maintained in the list as shown using the different color code in following figures.

ਜਲੰਧਰ, 4 ਅਪ੍ਰੈਲ (ਐੱਮ. ਐੱਸ. ਲੋਹੀਆ)-ਨਿਊ ਦਿਓਲ ਨਗਰ 'ਚ ਬੰਦ ਪਏ ਮਕਾਨ 'ਚੋਂ ਕਿਸੇ ਨੇ ਨਕਦੀ, ਗਹਿਣੇ ਤੇ ਹੋਰ ਸਾਮਾਨ ਚੋਰੀ ਕਰ ਲਿਆ। ਚੋਰੀ ਦੀ ਵਾਰਦਾਤ ਦੇ ਸ਼ਿਕਾਰ ਦੀਪਕ ਕਾਲੜਾ ਨੇ ਦੱਸਿਆ ਕਿ ਉਹ ਫ਼ਾਸਟ ਫੂਡ ਦਾ ਕੰਮ ਕਰਦਾ ਹੈ। ਉਹ ਸ਼ੁਕਰਵਾਰ ਨੂੰ ਆਪਣੇ ਪਰਿਵਾਰ ਨਾਲ ਇਕ ਧਾਰਮਿਕ ਸਥਾਨ 'ਤੇ ਗਿਆ ਸੀ, ਜਿਸ ਕਰਕੇ ਘਰ ਨੂੰ ਤਾਲੇ ਲਗਾਏ ਹੋਏ ਸਨ। ਜਦੋਂ ਅੱਜ ਉਹ ਘਰ ਵਾਪਸ ਪਰਤੇ ਤਾਂ ਮੁੱਖ ਦਰਵਾਜ਼ੇ ਦੇ ਤਾਲੇ ਤਾਂ ਲੱਗੇ ਹੋਏ ਸਨ, ਪਰ ਅੰਦਰਲੇ ਕਮਰਿਆਂ ਦੇ ਤਾਲੇ ਟੁੱਟੇ ਹੋਏ ਸਨ। ਅੰਦਰ ਜਾਂਚ ਕਰਨ 'ਤੇ ਪਤਾ ਲੱਗਾ ਕਿ ਘਰ 'ਚ ਪਈ 15 ਹਜ਼ਾਰ ਦੀ ਨਕਦੀ, 2 ਤੋਲੇ ਸੋਨੇ ਦੇ ਗਹਿਣੇ, 2 ਸਿਲੰਡਰ ਅਤੇ ਚੜ੍ਹਾਏ ਵਾਲੀ ਗੈਲਕ ਚੋਰੀ ਹੋ ਗਈ ਹੈ। ਇਸ ਸਬੰਧੀ ਪੁਲਿਸ ਨੂੰ ਸੂਚਨਾ ਦੇ ਦਿੱਤੀ ਹੈ।

ਲੁਧਿਆਣਾ, 2 ਅਪ੍ਰੈਲ (ਅਮਰੀਕ ਸਿੰਘ ਬੱਤਰਾ)- ਰਾਜ ਸਰਕਾਰ ਵੱਲੋਂ ਥਰਮੋਕੋਲ (ਡਿਸਪੋਜ਼ਏਬਲ) ਸਮਾਨ ਅਤੇ ਪਲਾਸਟਿਕ ਦੇ ਕੈਰੀਬੈਗਜ਼ ਦੇ ਉਤਪਾਦਨ ਅਤੇ ਵਿਕਰੀ ਤੇ ਲਗਾਈ ਰੋਕ ਨੂੰ ਲਾਗੂ ਕਰਾਉਣ ਲਈ ਪਹਿਲੀ ਅਪ੍ਰੈਲ ਤੋਂ ਨਗਰ ਨਿਗਮ ਪ੍ਰਸਾਸਨ ਨੇ ਸ਼ੁਰੂ ਕੀਤੀ ਕਾਰਵਾਈ ਦਾ ਪਲਾਸਟਿਕ ਮੈਨੂਫ਼ੈਕਚਰਜ਼ ਐਸੋਸੀਏਸਨ ਪੰਜਾਬ ਅਤੇ ਆਮ ਦੁਕਾਨਦਾਰਾਂ ਵੱਲੋਂ ਸਖਤ ਵਿਰੋਧ ਕੀਤਾ ਜਾ ਰਿਹਾ ਹੈ ਜਿਸ ਕਾਰਨ ਨਗਰ ਨਿਗਮ ਸਟਾਫ਼ ਅਤੇ ਦੁਕਾਨਦਾਰਾਂ ਦਰਮਿਆਨ ਬਣ ਰਹੀ ਤਣਾਅ ਦੀ ਸਥਿਤੀ ਨੂੰ ਸਾਂਤ ਕਰਨ ਲਈ ਮੇਅਰ ਹਰਚਰਨ ਸਿੰਘ ਗੋਹਲਵੜੀਆ ਨੇ ਨਗਰ ਨਿਗਮ ਸਿਹਤ ਸਾਖਾ ਅਧਿਕਾਰੀਆਂ ਅਤੇ ਦੁਕਾਨਦਾਰਾਂ ਦਰਮਿਆਨ ਮੀਟਿੰਗ ਸੱਦਣ ਦੀ ਹਦਾਇਤ ਦਿੱਤੀ ਹੈ। ਨਗਰ ਨਿਗਮ ਸਿਹਤ ਸਾਖਾ ਅਧਿਕਾਰੀਆਂ/ਕਰਮਚਾਰੀਆਂ ਵੱਲੋਂ ਸਹਿਰ ਦੇ ਹਲਵਾਈਆਂ, ਬੇਕਰੀ ਸ਼ਾਪ, ਕਰਿਆਨਾ ਅਤੇ ਖਾਣ ਪੀਣ ਦਾ ਦੂਸਰਾ ਸਮਾਨ ਵੇਚਣ ਵਾਲੇ ਦੁਕਾਨਦਾਰਾਂ ਨੂੰ ਪਲਾਸਟਿਕ ਕੈਰੀ ਬੈਗਜ਼ ਤੇ ਥਰਮੋਕੋਲ ਦਾ ਸਮਾਨ ਨਾ ਵੇਚਣ ਲਈ ਜਾਗਰੂਕ ਕਰਨ ਵਾਸਤੇ ਸ਼ੁਰੂ ਕੀਤੀ ਕਾਰਵਾਈ ਦੌਰਾਨ ਸ਼ੁੱਕਰਵਾਰ ਸਾਮ ਨੂੰ ਮਿਊਾਸਪਲ ਸੈਨੇਟਰੀ ਅਫ਼ਸਰ ਅਮੀਰ ਸਿੰਘ ਬਾਜਵਾ ਦੀ ਅਗਵਾਈ ਹੇਠ ਸਿਹਤ ਸਾਖਾ ਸਟਾਫ਼ ਵੱਲੋਂ ਸਿੰਗਾਰ ਸਿਨੇਮਾ ਰੋਡ 'ਤੇ ਦੁਕਾਨਾਂ ਦੀ ਕੀਤੀ ਜਾ ਰਹੀ ਚੈਕਿੰਗ ਦੌਰਾਨ ਅਧਿਕਾਰੀਆਂ ਦੀ ਦੁਕਾਨਦਾਰ ਨਾਲ ਤਿੱਖੀ ਬਹਿਸ ਹੋ ਗਈ ਸੀ। ਦੋਵੇਂ ਧਿਰਾਂ ਵੱਲੋਂ ਜ਼ੋਨਲ ਕਮਿਸ਼ਨਰ ਕਮਲੇਸ਼ ਬਾਂਸਲ ਨੂੰ ਦਿੱਤੀ ਸਿਕਾਇਤ ਵਿਚ ਇਕ ਦੂਜੇ 'ਤੇ ਦੁਰਵਿਵਹਾਰ, ਮੰਦੀ ਭਾਸਾ ਬੋਲਣ ਅਤੇ ਧੱਕਾਮੁੱਕੀ ਦਾ ਦੋਸ ਲਗਾਇਆ ਸੀ। ਜ਼ੋਨਲ ਕਮਿਸਨਰ ਨੇ ਦੋਵਾਂ ਧਿਰਾਂ ਖਿਲਾਫ਼ ਬਣਦੀ ਕਾਰਵਾਈ ਲਈ ਰਿਪੋਰਟ ਕਮਿਸਨਰ ਨੂੰ ਭੇਜੀ ਦੱਸੀ ਜਾ ਰਹੀ ਹੈ। ਦੁਕਾਨਦਾਰਾਂ ਦੇ ਹੱਕ ਵਿਚ ਨਿੱਤਰੇ ਭਾਜਪਾ ਕੌਂਸਲਰ ਇੰਦਰ ਅਗਰਵਾਲ ਤੋਂ ਬਾਅਦ ਮੌਕੇ 'ਤੇ ਪੁੱਜੇ ਜਿਲ੍ਹਾ ਭਾਜਪਾ ਪ੍ਰਧਾਨ ਪ੍ਰਵੀਨ ਬਾਂਸਲ, ਡਾਕਟਰ ਸੁਭਾਸ ਵਰਮਾ, ਭਾਜਪਾ ਕੌਂਸਲਰ ਦਲ ਦੇ ਆਗੂ ਗੁਰਦੀਪ ਸਿੰਘ ਨੀਟੂ, ਭਾਜਪਾ ਜ਼ਿਲ੍ਹਾ ਗਰੁਪ ਪ੍ਰਧਾਨ ਸਰਬਜੀਤ ਸਿੰਘ ਕਾਕਾ ਨੇ ਮੇਅਰ ਹਰਚਰਨ ਸਿੰਘ ਗੋਹਲਵੜੀਆ ਨੂੰ ਮਿਲਕੇ ਸਥਿਤੀ ਜਾਣੂ ਕਰਾਕੇ ਲੁੜੀਂਦੀ ਕਾਰਵਾਈ ਦੀ ਮੰਗ ਕੀਤੀ ਸੀ। ਜਿਸ 'ਤੇ ਮੇਅਰ ਗੋਹਲਵੜੀਆ ਨੇ ਨਗਰ ਨਿਗਮ ਅਧਿਕਾਰੀਆਂ ਨੂੰ ਹਦਾਇਤ ਕੀਤੀ ਕਿ ਸੋਮਵਾਰ ਨੂੰ ਦੁਕਾਨਦਾਰਾਂ ਨਾਲ ਮੀਟਿੰਗ ਕਰਕੇ ਕੀਤੀ ਜਾਣ ਵਾਲੀ ਕਾਰਵਾਈ ਤੋਂ ਜਾਣੂ ਕਰਾਉਣ ਤੋਂ ਬਾਅਦ ਪਾਬੰਦੀਸੁਦਾ ਸਮਾਨ ਵਿਰੁੱਧ ਕਾਰਵਾਈ ਕਰਨ/ਜ਼ਬਤ ਕਰਨ ਦੀ ਕਾਰਵਾਈ ਸ਼ੁਰੂ ਕੀਤੀ ਜਾਵੇਗੀ। ਇਸ ਦੌਰਾਨ ਪ੍ਰਧਾਨ ਜਥੇਦਾਰ ਹਰਭਜਨ ਸਿੰਘ ਡੰਗ ਦੀ ਅਗਵਾਈ ਹੇਠ ਮੁੱਖ ਮੰਤਰੀ ਪੰਜਾਬ ਦੇ ਸਿਆਸੀ ਸਲਾਹਕਾਰ ਮਹੇਸਇੰਦਰ ਸਿੰਘ ਗਰੇਵਾਲ ਨਾਲ ਮੁਲਾਕਾਤ ਕਰਕੇ ਪੇਸ ਆ ਰਹੀ ਮੁਸਕਿਲ ਤੋਂ ਨਿਜਾਤ ਦਿਵਾਉਣ ਦੀ ਮੰਗ ਕੀਤੀ ਜਿਸ 'ਤੇ ਗਰੇਵਾਲ ਵੱਲੋਂ ਉਚ ਅਧਿਕਾਰੀਆਂ ਨਾਲ ਪਲਾਸਟਿਕ ਕਾਰੋਬਾਰੀਆਂ ਦੀ ਮੰਗਲਵਾਰ ਸਵੇਰੇ 11 ਵਜੇ ਮੀਟਿੰਗ ਤਹਿ ਕਰਾਈ ਹੈ। ਮੀਟਿੰਗ ਵਿਚ ਹੋਣ ਵਾਲੇ ਫ਼ੈਸਲੇ ਅਨੁਸਾਰ ਅਗਲੇਰੀ ਕਾਰਵਾਈ ਕੀਤੀ ਜਾਵੇਗੀ।

3

Figure 4.3 Sentence in Red Color Representing First Block and Contents of List

ਜਲੰਧਰ, 4 ਅਪ੍ਰੈਲ (ਐੱਮ. ਐੱਸ. ਲੋਹੀਆ)-ਨਿਊ ਦਿਓਲ ਨਗਰ 'ਚ ਬੰਦ ਪਏ ਮਕਾਨ 'ਚੋਂ ਕਿਸੇ ਨੇ ਨਕਦੀ, ਗਹਿਣੇ ਤੇ ਹੋਰ ਸਾਮਾਨ ਚੋਰੀ ਕਰ ਲਿਆ। ਚੋਰੀ ਦੀ ਵਾਰਦਾਤ ਦੇ ਸ਼ਿਕਾਰ ਦੀਪਕ ਕਾਲੜਾ ਨੇ ਦੱਸਿਆ ਕਿ ਉਹ ਫ਼ਾਸਟ ਫੂਡ ਦਾ ਕੰਮ ਕਰਦਾ ਹੈ। ਉਹ ਸ਼ੁਕਰਵਾਰ ਨੂੰ ਆਪਣੇ ਪਰਿਵਾਰ ਨਾਲ ਇਕ ਧਾਰਮਿਕ ਸਥਾਨ 'ਤੇ ਗਿਆ ਸੀ, ਜਿਸ ਕਰਕੇ ਘਰ ਨੂੰ ਤਾਲੇ ਲਗਾਏ ਹੋਏ ਸਨ। ਜਦੋਂ ਅੱਜ ਉਹ ਘਰ ਵਾਪਸ ਪਰਤੇ ਤਾਂ ਮੁੱਖ ਦਰਵਾਜ਼ੇ ਦੇ ਤਾਲੇ ਤਾਂ ਲੱਗੇ ਹੋਏ ਸਨ, ਪਰ ਅੰਦਰਲੇ ਕਮਰਿਆਂ ਦੇ ਤਾਲੇ ਟੁੱਟੇ ਹੋਏ ਸਨ। ਅੰਦਰ ਜਾਂਚ ਕਰਨ 'ਤੇ ਪਤਾ ਲੱਗਾ ਕਿ ਘਰ 'ਚ ਪਈ 15 ਹਜ਼ਾਰ ਦੀ ਨਕਦੀ, 2 ਤੋਲੇ ਸੋਨੇ ਦੇ ਗਹਿਣੇ, 2 ਸਿਲੰਡਰ ਅਤੇ ਚੜ੍ਹਾਏ ਵਾਲੀ ਗੋਲਕ ਚੋਰੀ ਹੋ ਗਈ ਹੈ। ਇਸ ਸਬੰਧੀ ਪੁਲਿਸ ਨੂੰ ਸੂਚਨਾ ਦੇ ਦਿੱਤੀ ਹੈ।

3
10

ਲੁਧਿਆਣਾ, 2 ਅਪ੍ਰੈਲ (ਅਮਰੀਕ ਸਿੰਘ ਬੱਤਰਾ)- ਰਾਜ ਸਰਕਾਰ ਵੱਲੋਂ ਥਰਮੋਕੋਲ (ਡਿਸਪੋਜ਼ਏਬਲ) ਸਮਾਨ ਅਤੇ ਪਲਾਸਟਿਕ ਦੇ ਕੈਰੀਬੈਗਜ਼ ਦੇ ਉਤਪਾਦਨ ਅਤੇ ਵਿਕਰੀ ਤੇ ਲਗਾਈ ਰੋਕ ਨੂੰ ਲਾਗੂ ਕਰਾਉਣ ਲਈ ਪਹਿਲੀ ਅਪ੍ਰੈਲ ਤੋਂ ਨਗਰ ਨਿਗਮ ਪ੍ਰਸ਼ਾਸਨ ਨੇ ਸ਼ੁਰੂ ਕੀਤੀ ਕਾਰਵਾਈ ਦਾ ਪਲਾਸਟਿਕ ਮੈਨੂਫ਼ੈਕਚਰਜ਼ ਐਸੋਸੀਏਸ਼ਨ ਪੰਜਾਬ ਅਤੇ ਆਮ ਦੁਕਾਨਦਾਰਾਂ ਵੱਲੋਂ ਸਖਤ ਵਿਰੋਧ ਕੀਤਾ ਜਾ ਰਿਹਾ ਹੈ ਜਿਸ ਕਾਰਨ ਨਗਰ ਨਿਗਮ ਸਟਾਫ਼ ਅਤੇ ਦੁਕਾਨਦਾਰਾਂ ਦਰਮਿਆਨ ਬਣ ਰਹੀ ਤਣਾਅ ਦੀ ਸਥਿਤੀ ਨੂੰ ਸ਼ਾਂਤ ਕਰਨ ਲਈ ਮੇਅਰ ਹਰਚਰਨ ਸਿੰਘ ਗੋਹਲਵੜੀਆ ਨੇ ਨਗਰ ਨਿਗਮ ਸਿਹਤ ਸ਼ਾਖਾ ਅਧਿਕਾਰੀਆਂ ਅਤੇ ਦੁਕਾਨਦਾਰਾਂ ਦਰਮਿਆਨ ਮੀਟਿੰਗ ਸੱਦਣ ਦੀ ਹਦਾਇਤ ਦਿੱਤੀ ਹੈ। ਨਗਰ ਨਿਗਮ ਸਿਹਤ ਸ਼ਾਖਾ ਅਧਿਕਾਰੀਆਂ/ਕਰਮਚਾਰੀਆਂ ਵੱਲੋਂ ਸ਼ਹਿਰ ਦੇ ਹਲਵਾਈਆਂ, ਬੇਕਰੀ ਸ਼ਾਪ, ਕਰਿਆਨਾ ਅਤੇ ਖਾਣ ਪੀਣ ਦਾ ਦੂਸਰਾ ਸਮਾਨ ਵੇਚਣ ਵਾਲੇ ਦੁਕਾਨਦਾਰਾਂ ਨੂੰ ਪਲਾਸਟਿਕ ਕੈਰੀ ਬੈਗਜ਼ ਤੇ ਥਰਮੋਕੋਲ ਦਾ ਸਮਾਨ ਨਾ ਵੇਚਣ ਲਈ ਜਾਗਰੂਕ ਕਰਨ ਵਾਸਤੇ ਸ਼ੁਰੂ ਕੀਤੀ ਕਾਰਵਾਈ ਦੌਰਾਨ ਸ਼ੁੱਕਰਵਾਰ ਸ਼ਾਮ ਨੂੰ ਮਿਊਂਸਪਲ ਸੈਨੇਟਰੀ ਅਫ਼ਸਰ ਅਮੀਰ ਸਿੰਘ ਬਾਜਵਾ ਦੀ ਅਗਵਾਈ ਹੇਠ ਸਿਹਤ ਸਾਖਾ ਸਟਾਫ਼ ਵੱਲੋਂ ਸਿੰਗਾਰ ਸਿਨੇਮਾ ਰੋਡ 'ਤੇ ਦੁਕਾਨਾਂ ਦੀ ਕੀਤੀ ਜਾ ਰਹੀ ਚੈਕਿੰਗ ਦੌਰਾਨ ਅਧਿਕਾਰੀਆਂ ਦੀ ਦੁਕਾਨਦਾਰ ਨਾਲ ਤਿੱਖੀ ਬਹਿਸ ਹੋ ਗਈ ਸੀ। ਦੋਵੇਂ ਧਿਰਾਂ ਵੱਲੋਂ ਜ਼ੋਨਲ ਕਮਿਸ਼ਨਰ ਕਮਲੇਸ਼ ਬਾਂਸਲ ਨੂੰ ਦਿੱਤੀ ਸ਼ਿਕਾਇਤ ਵਿਚ ਇਕ ਦੂਜੇ 'ਤੇ ਦੁਰਵਿਵਹਾਰ, ਮੰਦੀ ਭਾਸ਼ਾ ਬੋਲਣ ਅਤੇ ਧੱਕਾਮੁੱਕੀ ਦਾ ਦੋਸ਼ ਲਗਾਇਆ ਸੀ। ਜ਼ੋਨਲ ਕਮਿਸ਼ਨਰ ਨੇ ਦੋਵਾਂ ਧਿਰਾਂ ਖਿਲਾਫ਼ ਬਣਦੀ ਕਾਰਵਾਈ ਲਈ ਰਿਪੋਰਟ ਕਮਿਸ਼ਨਰ ਨੂੰ ਭੇਜੀ ਦੱਸੀ ਜਾ ਰਹੀ ਹੈ। ਦੁਕਾਨਦਾਰਾਂ ਦੇ ਹੱਕ ਵਿਚ ਨਿੱਤਰੇ ਭਾਜਪਾ ਕੌਂਸਲਰ ਇੰਦਰ ਅਗਰਵਾਲ ਤੋਂ ਬਾਅਦ ਮੌਕੇ 'ਤੇ ਪੁੱਜੇ ਜ਼ਿਲ੍ਹਾ ਭਾਜਪਾ ਪ੍ਰਧਾਨ ਪ੍ਰਵੀਨ ਬਾਂਸਲ, ਡਾਕਟਰ ਸੁਭਾਸ਼ ਵਰਮਾ, ਭਾਜਪਾ ਕੌਂਸਲਰ ਦਲ ਦੇ ਆਗੂ ਗੁਰਦੀਪ ਸਿੰਘ ਨੀਟੂ, ਭਾਜਪਾ ਜ਼ਿਲ੍ਹਾ ਉਪ ਪ੍ਰਧਾਨ ਸਰਬਜੀਤ ਸਿੰਘ ਕਾਕਾ ਨੇ ਮੇਅਰ ਹਰਚਰਨ ਸਿੰਘ ਗੋਹਲਵੜੀਆ ਨੂੰ ਮਿਲਕੇ ਸਥਿਤੀ ਜਾਣੂ ਕਰਾਕੇ ਲੁੜੀਂਦੀ ਕਾਰਵਾਈ ਦੀ ਮੰਗ ਕੀਤੀ ਸੀ। ਜਿਸ 'ਤੇ ਮੇਅਰ ਗੋਹਲਵੜੀਆ ਨੇ ਨਗਰ ਨਿਗਮ ਅਧਿਕਾਰੀਆਂ ਨੂੰ ਹਦਾਇਤ ਕੀਤੀ ਕਿ ਸੋਮਵਾਰ ਨੂੰ ਦੁਕਾਨਦਾਰਾਂ ਨਾਲ ਮੀਟਿੰਗ ਕਰਕੇ ਕੀਤੀ ਜਾਣ ਵਾਲੀ ਕਾਰਵਾਈ ਤੋਂ ਜਾਣੂ ਕਰਾਉਣ ਤੋਂ ਬਾਅਦ ਪਾਬੰਦੀਸ਼ੁਦਾ ਸਮਾਨ ਵਿਰੁੱਧ ਕਾਰਵਾਈ ਕਰਨ/ਜ਼ਬਤ ਕਰਨ ਦੀ ਕਾਰਵਾਈ ਸ਼ੁਰੂ ਕੀਤੀ ਜਾਵੇਗੀ। ਇਸ ਦੌਰਾਨ ਪ੍ਰਧਾਨ ਜਥੇਦਾਰ ਹਰਭਜਨ ਸਿੰਘ ਡੰਗ ਦੀ ਅਗਵਾਈ ਹੇਠ ਮੁੱਖ ਮੰਤਰੀ ਪੰਜਾਬ ਦੇ ਸਿਆਸੀ ਸਲਾਹਕਾਰ ਮਹੇਸਇੰਦਰ ਸਿੰਘ ਗਰੇਵਾਲ ਨਾਲ ਮੁਲਾਕਾਤ ਕਰਕੇ ਪੇਸ਼ ਆ ਰਹੀ ਮੁਸ਼ਕਿਲ ਤੋਂ ਨਿਜਾਤ ਦਿਵਾਉਣ ਦੀ ਮੰਗ ਕੀਤੀ ਜਿਸ 'ਤੇ ਗਰੇਵਾਲ ਵੱਲੋਂ ਉਚ ਅਧਿਕਾਰੀਆਂ ਨਾਲ ਪਲਾਸਟਿਕ ਕਾਰੋਬਾਰੀਆਂ ਦੀ ਮੰਗਲਵਾਰ ਸਵੇਰੇ 11 ਵਜੇ ਮੀਟਿੰਗ ਤਹਿ ਕਰਾਈ ਹੈ। ਮੀਟਿੰਗ ਵਿਚ ਹੋਣ ਵਾਲੇ ਫ਼ੈਸਲੇ ਅਨੁਸਾਰ ਅਗਲੇਰੀ ਕਾਰਵਾਈ ਕੀਤੀ ਜਾਵੇਗੀ।

Figure 4.4: Sentence in Purple Color Representing Second Block and Contents of List

ਜਲੰਧਰ, 4 ਅਪ੍ਰੈਲ (ਐੱਮ. ਐੱਸ. ਲੋਹੀਆ)-ਨਿਊ ਦਿਓਲ ਨਗਰ 'ਚ ਬੰਦ ਪਏ ਮਕਾਨ 'ਚੋਂ ਕਿਸੇ ਨੇ ਨਕਦੀ, ਗਹਿਣੇ ਤੇ ਹੋਰ ਸਾਮਾਨ ਚੋਰੀ ਕਰ ਲਿਆ। ਚੋਰੀ ਦੀ ਵਾਰਦਾਤ ਦੇ ਸ਼ਿਕਾਰ ਦੀਪਕ ਕਾਲੜਾ ਨੇ ਦੱਸਿਆ ਕਿ ਉਹ ਫ਼ਾਸਟ ਫੂਡ ਦਾ ਕੰਮ ਕਰਦਾ ਹੈ। ਉਹ ਸ਼ੁਕਰਵਾਰ ਨੂੰ ਆਪਣੇ ਪਰਿਵਾਰ ਨਾਲ ਇਕ ਧਾਰਮਿਕ ਸਥਾਨ 'ਤੇ ਗਿਆ ਸੀ, ਜਿਸ ਕਰਕੇ ਘਰ ਨੂੰ ਤਾਲੇ ਲਗਾਏ ਹੋਏ ਸਨ। ਜਦੋਂ ਅੱਜ ਉਹ ਘਰ ਵਾਪਸ ਪਰਤੇ ਤਾਂ ਮੁੱਖ ਦਰਵਾਜ਼ੇ ਦੇ ਤਾਲੇ ਤਾਂ ਲੱਗੇ ਹੋਏ ਸਨ, ਪਰ ਅੰਦਰਲੇ ਕਮਰਿਆਂ ਦੇ ਤਾਲੇ ਟੁੱਟੇ ਹੋਏ ਸਨ। ਅੰਦਰ ਜਾਂਚ ਕਰਨ 'ਤੇ ਪਤਾ ਲੱਗਾ ਕਿ ਘਰ 'ਚ ਪਈ 15 ਹਜ਼ਾਰ ਦੀ ਨਕਦੀ, 2 ਤੋਲੇ ਸੋਨੇ ਦੇ ਗਹਿਣੇ, 2 ਸਿਲੰਡਰ ਅਤੇ ਚੜ੍ਹਾਏ ਵਾਲੀ ਗੋਲਕ ਚੋਰੀ ਹੋ ਗਈ ਹੈ। ਇਸ ਸਬੰਧੀ ਪੁਲਿਸ ਨੂੰ ਸੂਚਨਾ ਦੇ ਦਿੱਤੀ ਹੈ।

3
10
5

ਲੁਧਿਆਣਾ, 2 ਅਪ੍ਰੈਲ (ਅਮਰੀਕ ਸਿੰਘ ਬੱਤਰਾ)- ਰਾਜ ਸਰਕਾਰ ਵੱਲੋਂ ਥਰਮੋਕੋਲ (ਡਿਸਪੋਜ਼ਏਬਲ) ਸਮਾਨ ਅਤੇ ਪਲਾਸਟਿਕ ਦੇ ਕੈਰੀਬੈਗਜ਼ ਦੇ ਉਤਪਾਦਨ ਅਤੇ ਵਿਕਰੀ ਤੇ ਲਗਾਈ ਰੋਕ ਨੂੰ ਲਾਗੂ ਕਰਾਉਣ ਲਈ ਪਹਿਲੀ ਅਪ੍ਰੈਲ ਤੋਂ ਨਗਰ ਨਿਗਮ ਪ੍ਰਸ਼ਾਸਨ ਨੇ ਸ਼ੁਰੂ ਕੀਤੀ ਕਾਰਵਾਈ ਦਾ ਪਲਾਸਟਿਕ ਮੈਨੂਫ਼ੈਕਚਰਜ਼ ਐਸੋਸੀਏਸ਼ਨ ਪੰਜਾਬ ਅਤੇ ਆਮ ਦੁਕਾਨਦਾਰਾਂ ਵੱਲੋਂ ਸਖਤ ਵਿਰੋਧ ਕੀਤਾ ਜਾ ਰਿਹਾ ਹੈ ਜਿਸ ਕਾਰਨ ਨਗਰ ਨਿਗਮ ਸਟਾਫ਼ ਅਤੇ ਦੁਕਾਨਦਾਰਾਂ ਦਰਮਿਆਨ ਬਣ ਰਹੀ ਤਣਾਅ ਦੀ ਸਥਿਤੀ ਨੂੰ ਸ਼ਾਂਤ ਕਰਨ ਲਈ ਮੇਅਰ ਹਰਚਰਨ ਸਿੰਘ ਗੋਹਲਵੜੀਆ ਨੇ ਨਗਰ ਨਿਗਮ ਸਿਹਤ ਸ਼ਾਖਾ ਅਧਿਕਾਰੀਆਂ ਅਤੇ ਦੁਕਾਨਦਾਰਾਂ ਦਰਮਿਆਨ ਮੀਟਿੰਗ ਸੱਦਣ ਦੀ ਹਦਾਇਤ ਦਿੱਤੀ ਹੈ। ਨਗਰ ਨਿਗਮ ਸਿਹਤ ਸ਼ਾਖਾ ਅਧਿਕਾਰੀਆਂ/ਕਰਮਚਾਰੀਆਂ ਵੱਲੋਂ ਸ਼ਹਿਰ ਦੇ ਹਲਵਾਈਆਂ, ਬੇਕਰੀ ਸ਼ਾਪ, ਕਰਿਆਨਾ ਅਤੇ ਖਾਣ ਪੀਣ ਦਾ ਦੂਸਰਾ ਸਮਾਨ ਵੇਚਣ ਵਾਲੇ ਦੁਕਾਨਦਾਰਾਂ ਨੂੰ ਪਲਾਸਟਿਕ ਕੈਰੀ ਬੈਗਜ਼ ਤੇ ਥਰਮੋਕੋਲ ਦਾ ਸਮਾਨ ਨਾ ਵੇਚਣ ਲਈ ਜਾਗਰੂਕ ਕਰਨ ਵਾਸਤੇ ਸ਼ੁਰੂ ਕੀਤੀ ਕਾਰਵਾਈ ਦੌਰਾਨ ਸ਼ੁੱਕਰਵਾਰ ਸ਼ਾਮ ਨੂੰ ਮਿਊਂਸਪਲ ਸੈਨੇਟਰੀ ਅਫ਼ਸਰ ਅਮੀਰ ਸਿੰਘ ਬਾਜਵਾ ਦੀ ਅਗਵਾਈ ਹੇਠ ਸਿਹਤ ਸਾਖਾ ਸਟਾਫ਼ ਵੱਲੋਂ ਸਿੰਗਾਰ ਸਿਨੇਮਾ ਰੋਡ 'ਤੇ ਦੁਕਾਨਾਂ ਦੀ ਕੀਤੀ ਜਾ ਰਹੀ ਚੈਕਿੰਗ ਦੌਰਾਨ ਅਧਿਕਾਰੀਆਂ ਦੀ ਦੁਕਾਨਦਾਰ ਨਾਲ ਤਿੱਖੀ ਬਹਿਸ ਹੋ ਗਈ ਸੀ। ਦੋਵੇਂ ਧਿਰਾਂ ਵੱਲੋਂ ਜ਼ੋਨਲ ਕਮਿਸ਼ਨਰ ਕਮਲੇਸ਼ ਬਾਂਸਲ ਨੂੰ ਦਿੱਤੀ ਸ਼ਿਕਾਇਤ ਵਿਚ ਇਕ ਦੂਜੇ 'ਤੇ ਦੁਰਵਿਵਹਾਰ, ਮੰਦੀ ਭਾਸ਼ਾ ਬੋਲਣ ਅਤੇ ਧੱਕਾਮੁੱਕੀ ਦਾ ਦੋਸ਼ ਲਗਾਇਆ ਸੀ। ਜ਼ੋਨਲ ਕਮਿਸ਼ਨਰ ਨੇ ਦੋਵਾਂ ਧਿਰਾਂ ਖਿਲਾਫ਼ ਬਣਦੀ ਕਾਰਵਾਈ ਲਈ ਰਿਪੋਰਟ ਕਮਿਸ਼ਨਰ ਨੂੰ ਭੇਜੀ ਦੱਸੀ ਜਾ ਰਹੀ ਹੈ। ਦੁਕਾਨਦਾਰਾਂ ਦੇ ਹੱਕ ਵਿਚ ਨਿੱਤਰੇ ਭਾਜਪਾ ਕੌਂਸਲਰ ਇੰਦਰ ਅਗਰਵਾਲ ਤੋਂ ਬਾਅਦ ਮੌਕੇ 'ਤੇ ਪੁੱਜੇ ਜ਼ਿਲ੍ਹਾ ਭਾਜਪਾ ਪ੍ਰਧਾਨ ਪ੍ਰਵੀਨ ਬਾਂਸਲ, ਡਾਕਟਰ ਸੁਭਾਸ਼ ਵਰਮਾ, ਭਾਜਪਾ ਕੌਂਸਲਰ ਦਲ ਦੇ ਆਗੂ ਗੁਰਦੀਪ ਸਿੰਘ ਨੀਟੂ, ਭਾਜਪਾ ਜ਼ਿਲ੍ਹਾ ਉਪ ਪ੍ਰਧਾਨ ਸਰਬਜੀਤ ਸਿੰਘ ਕਾਕਾ ਨੇ ਮੇਅਰ ਹਰਚਰਨ ਸਿੰਘ ਗੋਹਲਵੜੀਆ ਨੂੰ ਮਿਲਕੇ ਸਥਿਤੀ ਜਾਣੂ ਕਰਾਕੇ ਲੁੜੀਂਦੀ ਕਾਰਵਾਈ ਦੀ ਮੰਗ ਕੀਤੀ ਸੀ। ਜਿਸ 'ਤੇ ਮੇਅਰ ਗੋਹਲਵੜੀਆ ਨੇ ਨਗਰ ਨਿਗਮ ਅਧਿਕਾਰੀਆਂ ਨੂੰ ਹਦਾਇਤ ਕੀਤੀ ਕਿ ਸੋਮਵਾਰ ਨੂੰ ਦੁਕਾਨਦਾਰਾਂ ਨਾਲ ਮੀਟਿੰਗ ਕਰਕੇ ਕੀਤੀ ਜਾਣ ਵਾਲੀ ਕਾਰਵਾਈ ਤੋਂ ਜਾਣੂ ਕਰਾਉਣ ਤੋਂ ਬਾਅਦ ਪਾਬੰਦੀਸ਼ੁਦਾ ਸਮਾਨ ਵਿਰੁੱਧ ਕਾਰਵਾਈ ਕਰਨ/ਜ਼ਬਤ ਕਰਨ ਦੀ ਕਾਰਵਾਈ ਸ਼ੁਰੂ ਕੀਤੀ ਜਾਵੇਗੀ। ਇਸ ਦੌਰਾਨ ਪ੍ਰਧਾਨ ਜਥੇਦਾਰ ਹਰਭਜਨ ਸਿੰਘ ਡੰਗ ਦੀ ਅਗਵਾਈ ਹੇਠ ਮੁੱਖ ਮੰਤਰੀ ਪੰਜਾਬ ਦੇ ਸਿਆਸੀ ਸਲਾਹਕਾਰ ਮਹੇਸਇੰਦਰ ਸਿੰਘ ਗਰੇਵਾਲ ਨਾਲ ਮੁਲਾਕਾਤ ਕਰਕੇ ਪੇਸ਼ ਆ ਰਹੀ ਮੁਸਕਿਲ ਤੋਂ ਨਿਜਾਤ ਦਿਵਾਉਣ ਦੀ ਮੰਗ ਕੀਤੀ ਜਿਸ 'ਤੇ ਗਰੇਵਾਲ ਵੱਲੋਂ ਉਚ ਅਧਿਕਾਰੀਆਂ ਨਾਲ ਪਲਾਸਟਿਕ ਕਾਰੋਬਾਰੀਆਂ ਦੀ ਮੰਗਲਵਾਰ ਸਵੇਰੇ 11 ਵਜੇ ਮੀਟਿੰਗ ਤਹਿ ਕਰਾਈ ਹੈ। ਮੀਟਿੰਗ ਵਿਚ ਹੋਣ ਵਾਲੇ ਫ਼ੈਸਲੇ ਅਨੁਸਾਰ ਅਗਲੇਰੀ ਕਾਰਵਾਈ ਕੀਤੀ ਜਾਵੇਗੀ।

Figure 4.5: Sentence in Ocher Color Representing Third Block and Contents of List

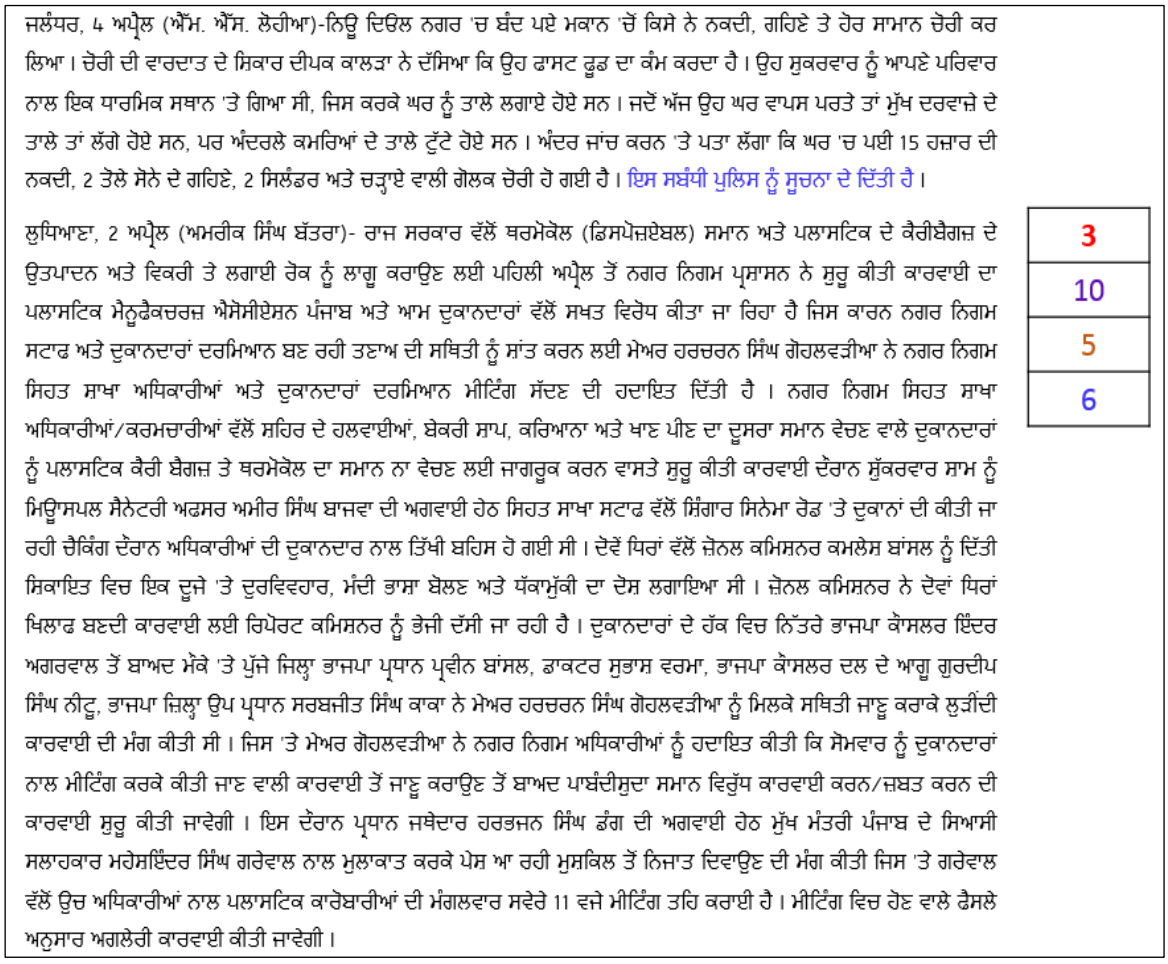

Figure 4.6: Sentence in Blue Color Representing Fourth Block and Contents of List

C. Output

The generated output of first layer embedding is the contents shown in Table 4.1 below.

Table 4.1: Output of First Layer Hiding

3	10	5	6

4.2.1.2 Second Layer Embedding

A. Input

The output of the first layer i.e. 3, 10, 5, 6 and same cover text file which has been used in first layer embedding as shown in Fig. 4.2.

B. Intermediate Steps

Step 1: The output of the first layer i.e. 3, 10, 5, 6 is converted into binary form. The binary form is **00000011000010100000010100000110**.

Step 3: scan the cover text file “cover.txt” to find the space between two words for embedding bits calculated in step 1. If space is found, LRM character is inserted for

bit 0 and RLM character is inserted for bit 1. In this way, all bits are embedded in the cover text file. Bits embedded in the cover text file is shown using the color code in following Fig. 4.7 where red color represents LRM and Green color for RLM character. A red color arrow indicates the insertion of LRM character at that position and a green arrow indicates the insertion of RLM character at that position.

Figure 4.7: Red Color Arrow Indicating Insertion of LMR and Green Color for RLM

C. Output

The generated output of second layer embedding is the final stego-Text at sender side shown in Fig. 4.8.

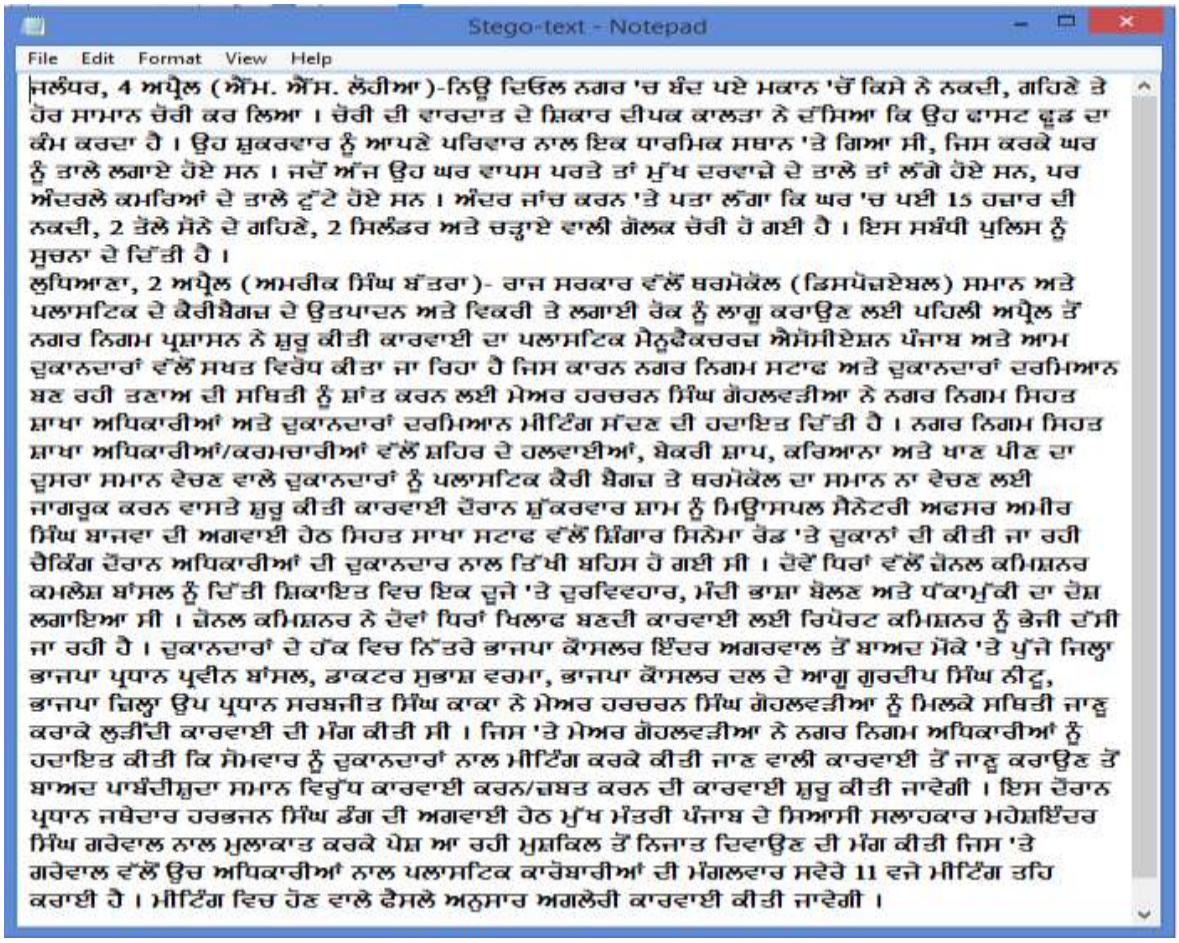

Figure 4.8: Final Generated Stego-Text File.

4.2.2 Extraction Process

The following figure shows the graphical user interface at receiver side for the extraction process.

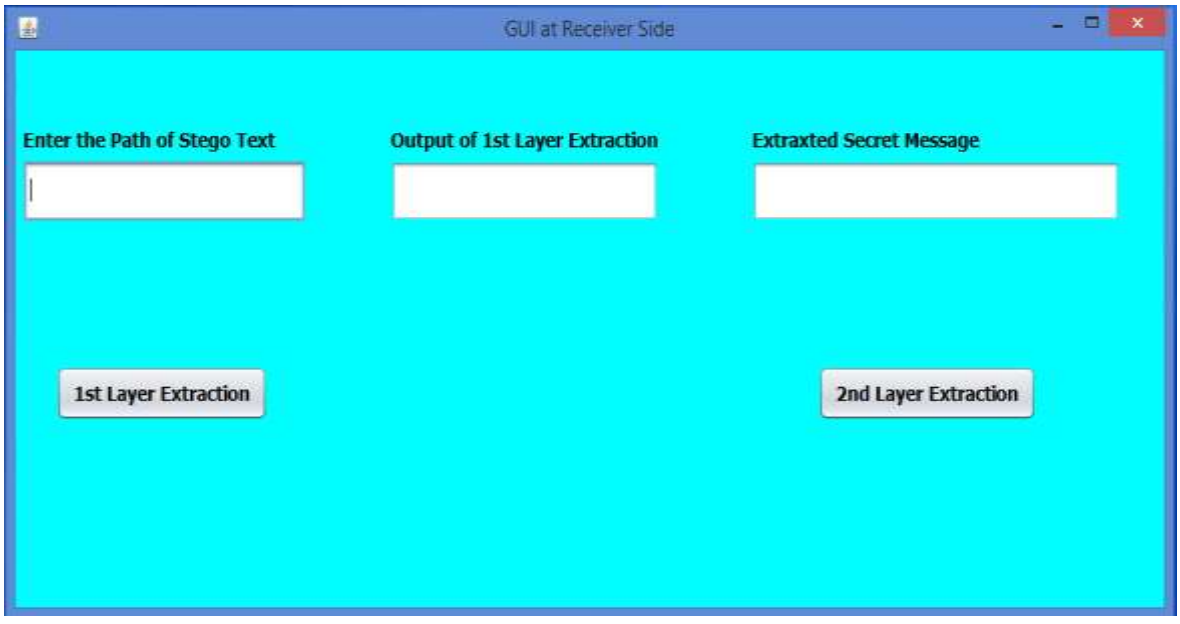

Figure 4.9: GUI for Receiver

4.2.2.1 First Layer Extraction

A. Input

The stego-Text file named “stego-text.txt” generated through embedding process shown in Fig.4.8.

B. Intermediate Steps

Step 1: scan the Stego-Text file “stego-text.txt” to find the LRM and RLM characters. If character found is LMR, it represents **‘0’**. If character found is RLM, it represents **‘1’**.

Step 2: According to the character found, a binary string is created. The binary string created is **00000011000010100000010100000110.**

Step 3: For each octet in binary string, a decimal value is evaluated. The decimal values are 3, 10, 5 and 6.

C. Output

The decimal values calculated at step 3 is the output of the first layer extraction process which indicates the sentence’s number in stago-text.

4.2.2.2 Second Layer Extraction

A. Input

The output of the first layer i.e. 3, 10, 5, 6 and same Stego-Text file which has been used in first layer extraction shown in Fig. 4.8.

B. Intermediate Steps

Step 1: Scan the Stego-Text file to locate the sentence number 3, 10, 5 and then 6.

Step 2: Find the properties of these sentences and represents 4-bit nibble according to properties found. The 4-bits nibble for sentence number 3, 10, 5 and 6 are 0101, 0100, 0110 and 1111 respectively.

Step 3: Concatenate all 4-bit blocks i.e. **0101010001101111**

Step 4: Get the alphabet for each octet i.e. **‘T’** and **‘o’**

C. Output

The hidden message is **‘To’**.

4.3 Similarity Measure Analysis

Similarity measures are essential to solve many pattern recognition problems such as classification, clustering, and retrieval problem. Similarity metrics are used to measure the distance between two values or sequences.

4.2.3 Jaro-Winkler Similarity Metric

The similarity between two strings can be calculated by using Jaro-Winkler metric. (Bhattacharyya, Banerjee & Sanyal, 2010). The Jaro-Winkler metric is applicable for measuring the similarity between the cover /file and stego file generated by using the proposed two layered hiding approach because it deals with strings. The Jaro-Winkler metric is calculated as**,**

$$\mathbf{jaro-Winkler(S1,S2) = jaro(S1,S2) + (L * P * (1 - jaro(S1,S2))}$$

$$\mathbf{jaro(S1,S2) = \frac{1}{3} * \left(\frac{m}{|S1|} + \frac{m}{|S2|} + \frac{m-t}{m}\right)}$$

Where S1 and S2 are the two strings whose similarity is to be measured, L is the length of common prefix (maximum 4 characters), P is scaling factor whose standard value is 0.1, m is the number of matching characters and t is the number of transpositions.

A high Jaro score represents greater similarity between the strings. Jaro score 0 indicates that the strings are not similar and 1 represents that they are exactly same. This analysis consists of two approaches .The first approach uses a fixed length (bits) of secret message and different lengths (bits) of cover text, whereas the second approach involves a fixed length (bits) of cover text and different lengths (bits) of secret messages. The details of both approaches are described below.

4.3.1.1 First Approach: Fix Secret Message with Different Cover Text

Table 4.2: Secret Message, Different Cover Files and Corresponding Stego-Text Files.

SECRET MESSAGE	COVER FILE	STEGO FILE
ABC	Cover1.txt	Stego1.txt
	Cover2.txt	Stego2.txt
	Cover3.txt	Stego3.txt
	Cover4.txt	Stego4.txt
	Cover5.txt	Stego5.txt

Following Tables show the Jaro score of the string pairs when the Jaro-Winkler metric is applied to the cover file (C) and the stego file (S). Where C_i is the i^{th} sentence of the cover file (C) and S_i is the corresponding i^{th} sentence the stego file (S).

Table 4.3: Jaro Score of the String Pairs of the Cover1.txt and Stego1.txt

String Pair	Jaro score
C_1S_1	0.960784
C_2S_2	0.960784
C_3S_3	1.0
C_4S_4	1.0
C_5S_5	1.0
C_6S_6	1.0
C_7S_7	1.0
C_8S_8	1.0
C_9S_9	1.0
$C_{10}S_{10}$	1.0
$C_{11}S_{11}$	1.0
$C_{12}S_{12}$	1.0
$C_{13}S_{13}$	1.0
$C_{14}S_{14}$	1.0
$C_{15}S_{15}$	1.0
Average Jaro Score=0.993905	

Table 4.4: Jaro Score of the String Pairs of the Cover2.txt and Stego2.txt

String Pair	Jaro score
C_1S_1	0.9602
C_2S_2	0.9976
C_3S_3	0.9975
C_4S_4	0.9982
C_5S_5	0.9981
C_6S_6	0.9985
C_7S_7	0.9983
C_8S_8	1.0
C_9S_9	1.0
$C_{10}S_{10}$	1.0
$C_{11}S_{11}$	1.0

(Cntd..)

Table 4.4 (Cntd..)

$C_{12}S_{12}$	1.0
$C_{13}S_{13}$	1.0
$C_{14}S_{14}$	1.0
$C_{15}S_{15}$	1.0
Average Jaro Score=0.996592	

Table 4.5: Jaro Score of the String Pairs of the Cover3.txt and Stego3.txt

String Pair	Jaro score
C_1S_1	0.952518
C_2S_2	0.990792
C_3S_3	1.0
C_4S_4	1.0
C_5S_5	1.0
C_6S_6	1.0
C_7S_7	1.0
C_8S_8	1.0
C_9S_9	1.0
$C_{10}S_{10}$	1.0
$C_{11}S_{11}$	1.0
$C_{12}S_{12}$	1.0
$C_{13}S_{13}$	1.0
$C_{14}S_{14}$	1.0
$C_{15}S_{15}$	1.0
$C_{16}S_{16}$	1.0
$C_{17}S_{17}$	1.0
$C_{18}S_{18}$	1.0
Average Jaro Score=0.996851	

Table 4.6: Jaro Score of the String Pairs of the Cover4.txt and Stego4.txt

String Pair	Jaro score
C_1S_1	0.951316
C_2S_2	0.977226
C_3S_3	1.0
C_4S_4	1.0
C_5S_5	1.0
C_6S_6	1.0
C_7S_7	1.0
C_8S_8	1.0
C_9S_9	1.0
$C_{10}S_{10}$	1.0
$C_{11}S_{11}$	1.0
$C_{12}S_{12}$	1.0
$C_{13}S_{13}$	1.0
$C_{14}S_{14}$	1.0
$C_{15}S_{15}$	1.0
$C_{16}S_{16}$	1.0
$C_{17}S_{17}$	1.0
$C_{18}S_{18}$	1.0
$C_{19}S_{19}$	1.0
$C_{20}S_{20}$	1.0
Average Jaro Score=0.996427	

Table 4.7: Jaro Score of the String Pairs of the Cover5.txt and Stego5.txt

String Pair	Jaro score
C_1S_1	0.951316
C_2S_2	0.977226
C_3S_3	1.0
C_4S_4	1.0
C_5S_5	1.0
C_6S_6	1.0

(Cntd..)

Table 4.7 (Cntd..)

C_7S_7	1.0
C_8S_8	1.0
C_9S_9	1.0
$C_{10}S_{10}$	1.0
$C_{11}S_{11}$	1.0
$C_{12}S_{12}$	1.0
$C_{13}S_{13}$	1.0
$C_{14}S_{14}$	1.0
$C_{15}S_{15}$	1.0
$C_{16}S_{16}$	1.0
$C_{17}S_{17}$	1.0
$C_{18}S_{18}$	1.0
$C_{19}S_{19}$	1.0
$C_{20}S_{20}$	1.0
$C_{21}S_{21}$	1.0
Average Jaro Score=0.995068	

Table 4.8 shows average Jaro Score of five cover files and their corresponding stego files calculated above when the Jaro-Winkler metric is applied.

Table 4.8: Average Jaro Score of Five Samples for First Approach

Sample	I	II	III	IV	V	Average Jaro Score=0.995769
Jaro score	0.993905	0.996592	0.996851	0.996427	0.995068	

4.3.1.2 Second Approach: Fix Cover Text with Different Secret Messages

Table 4.9: Different Secret Messages, Cover file and Corresponding Stego-Text Files.

SECRET MESSAGE	COVER FILE	STEGO FILE
M	Cover.txt	Stego_M.txt
DO		Stego_DO.txt
ARE		Stego_ARE.txt
foot		Stego_foot.txt
cover		Stego_cover.txt

Following Tables show the Jaro score of the string pairs when the Jaro-Winkler metric is applied to the cover file (C) and the stego file (S). Where C_i is the i^{th} sentence of the cover file (C) and S_i is the corresponding i^{th} sentence the stego file (S).

Table 4.10 Jaro Score of the String Pairs of the Cover.txt and Stego_M.txt

String Pair	Jaro score
C_1S_1	0.966925
C_2S_2	1.0
C_3S_3	1.0
C_4S_4	1.0
C_5S_5	1.0
C_6S_6	1.0
C_7S_7	1.0
C_8S_8	1.0
C_9S_9	1.0
$C_{10}S_{10}$	1.0
$C_{11}S_{11}$	1.0
$C_{12}S_{12}$	1.0
$C_{13}S_{13}$	1.0
$C_{14}S_{14}$	1.0
$C_{15}S_{15}$	1.0
$C_{16}S_{16}$	1.0
$C_{17}S_{17}$	1.0
$C_{18}S_{18}$	1.0
$C_{19}S_{19}$	1.0
$C_{20}S_{20}$	1.0
Average Jaro Score=0.996427	

Table 4.11: Jaro Score of the String Pairs of the Cover.txt and Stego_DO.txt

String Pair	Jaro score
C_1S_1	0.954902
C_2S_2	0.965116
C_3S_3	1.0
C_4S_4	1.0
C_5S_5	1.0
C_6S_6	1.0
C_7S_7	1.0
C_8S_8	1.0
C_9S_9	1.0
$C_{10}S_{10}$	1.0
$C_{11}S_{11}$	1.0
$C_{12}S_{12}$	1.0
$C_{13}S_{13}$	1.0
$C_{14}S_{14}$	1.0
$C_{15}S_{15}$	1.0
$C_{16}S_{16}$	1.0
$C_{17}S_{17}$	1.0
$C_{18}S_{18}$	1.0
$C_{19}S_{19}$	1.0
$C_{20}S_{20}$	1.0
Average Jaro Score=0.996001	

Table 4.12: Jaro Score of the String Pairs of the Cover.txt and Stego_ARE.txt

String Pair	Jaro score
C_1S_1	0.953285
C_2S_2	0.936842
C_3S_3	0.976898
C_4S_4	1.0
C_5S_5	1.0
C_6S_6	1.0

(Cntd..)

Table 4.12 (Cntd..)

C_7S_7	1.0
C_8S_8	1.0
C_9S_9	1.0
$C_{10}S_{10}$	1.0
$C_{11}S_{11}$	1.0
$C_{12}S_{12}$	1.0
$C_{13}S_{13}$	1.0
$C_{14}S_{14}$	1.0
$C_{15}S_{15}$	1.0
$C_{16}S_{16}$	1.0
$C_{17}S_{17}$	1.0
$C_{18}S_{18}$	1.0
$C_{19}S_{19}$	1.0
$C_{20}S_{20}$	1.0
Average Jaro Score=0.993351	

Table 4.13: Jaro Score of the String Pairs of the Cover.txt and stego_foot.txt

String Pair	Jaro score
C_1S_1	0.954902
C_2S_2	0.936842
C_3S_3	0.93913
C_4S_4	0.993884
C_5S_5	1.0
C_6S_6	1.0
C_7S_7	1.0
C_8S_8	1.0
C_9S_9	1.0
$C_{10}S_{10}$	1.0
$C_{11}S_{11}$	1.0
$C_{12}S_{12}$	1.0
$C_{13}S_{13}$	1.0

(Cntd..)

Table 4.13 (Cntd..)

$C_{14}S_{14}$	1.0
$C_{15}S_{15}$	1.0
$C_{16}S_{16}$	1.0
$C_{17}S_{17}$	1.0
$C_{18}S_{18}$	1.0
$C_{19}S_{19}$	1.0
$C_{20}S_{20}$	1.0
Average Jaro Score=0.991238	

Table 4.14: Jaro Score of the String Pairs of the cover.txt and stego_cover.txt

String Pair	Jaro score
C_1S_1	0.954902
C_2S_2	0.936842
C_3S_3	0.93913
C_4S_4	0.952
C_5S_5	1.0
C_6S_6	1.0
C_7S_7	1.0
C_8S_8	1.0
C_9S_9	1.0
$C_{10}S_{10}$	1.0
$C_{11}S_{11}$	1.0
$C_{12}S_{12}$	1.0
$C_{13}S_{13}$	1.0
$C_{14}S_{14}$	1.0
$C_{15}S_{15}$	1.0
$C_{16}S_{16}$	1.0
$C_{17}S_{17}$	1.0
$C_{18}S_{18}$	1.0
$C_{19}S_{19}$	1.0
$C_{20}S_{20}$	1.0
Average Jaro Score=0.989144	

Table 4.15 shows average Jaro Score of five cover files and their corresponding stego files calculated above when the Jaro-Winkler metric is applied.

Table 4.15: Average Jaro Score of Five Samples for the Second Approach

Sample	I	II	III	IV	V	Average Jaro Score=0.993616
Jaro score	0.998346	0.996001	0.993351	0.991238	0.989144	

Jaro score 0 indicates that the strings are not similar and 1 represents that they are exactly same. At second layer embedding in the proposed approach, LRM and RLM character are injected at space between words. Therefore, average jaro score calculated in Table 4.8 and Table 4.15 are 0.995769 and 0.993616 respectively, not equal to 1. The difference in these values compare to 1 is negligible and leading to the Jaro score of 1 which indicates that both the cover text and its corresponding stego file are exactly the same.

4.4 Capacity Ratio Analysis

Embedding Capacity also known as payload is the amount of data that can be hidden in a cover, compared to the size of the cover. A Steganographic algorithm with small embedding capacity may have other good features such as robustness. So it may be the ideal choice when only a small amount of data, such as a short message, has to be hidden. Normally capacity ratio is calculated as follows:

$$\boldsymbol{Capacity\ Ratio} = \frac{\boldsymbol{Amount\ of\ Hidden\ Bytes}}{\boldsymbol{Sizes\ of\ the\ cover\ text\ in\ Bytes}}$$

Assuming one character occupies one byte in memory, the percentage capacity is the capacity ratio multiplied by 100.

In proposed method, embedding is done at two levels. At first level, the attributes of a sentence in the cover text are specified by 4 bits. It means that there will be 16 options shown in Table 4.16. A sentence in cover text follows one of the 16 attributes. To hide a secret message i.e. a stream of binary digits of length N, there should be (N/4) numbers of sentences minimum in cover text. It means a single alphabet to be

hidden, there should be two lines minimum in cover file. A cover file of more than two sentences may still not be suitable to hide a single alphabet, if no sentence's attributes match to any 4 bits nibbles of given alphabet.

In case a cover media comprises of such sentences that specify all the above attributes can be used to hide message of any sizes. But at the second layer hiding, there should be enough amount of white space to accommodate the output of first layer hiding. Since the second layer embedding algorithm hides a single bit in space between two words. Therefore, at second layer embedding file space probability can be measured as follows:

$$\boldsymbol{Space\ Ratio} = \frac{\boldsymbol{Number\ of\ Spaces}}{\boldsymbol{Total\ Number\ of\ Bits}}$$

Table 4.16: All 4 Bits Combinations and their Corresponding Specification.

4-bits combinations	Specification of a sentence
0000	Vowel even, Consonant even, Characters even, Words even
0001	Vowel even, Consonant even, Characters even, Words odd
0010	Vowel even, Consonant even, Characters odd, Words even
0011	Vowel even, Consonant even, Characters odd, Words odd
0100	Vowel even, Consonant odd, Characters even, Words even
0101	Vowel even, Consonant odd, Characters even, Words odd
0110	Vowel even, Consonant odd, Characters odd, Words even
0111	Vowel even, Consonant odd, Characters odd, Words odd
1000	Vowel odd, Consonant even, Characters even, Words even
1001	Vowel odd, Consonant even, Characters even, Words odd
1010	Vowel odd, Consonant even, Characters odd, Words even
1011	Vowel odd, Consonant even, Characters odd, Words odd
1100	Vowel odd, Consonant odd, Characters even, Words even
1101	Vowel odd, Consonant odd, Characters even, Words odd
1110	Vowel odd, Consonant odd, Characters odd, Words even
1111	Vowel odd, Consonant odd, Characters odd, Words odd

Fig. 4.10 shows the relation between different file size and amount of data that can be hidden inside these files.

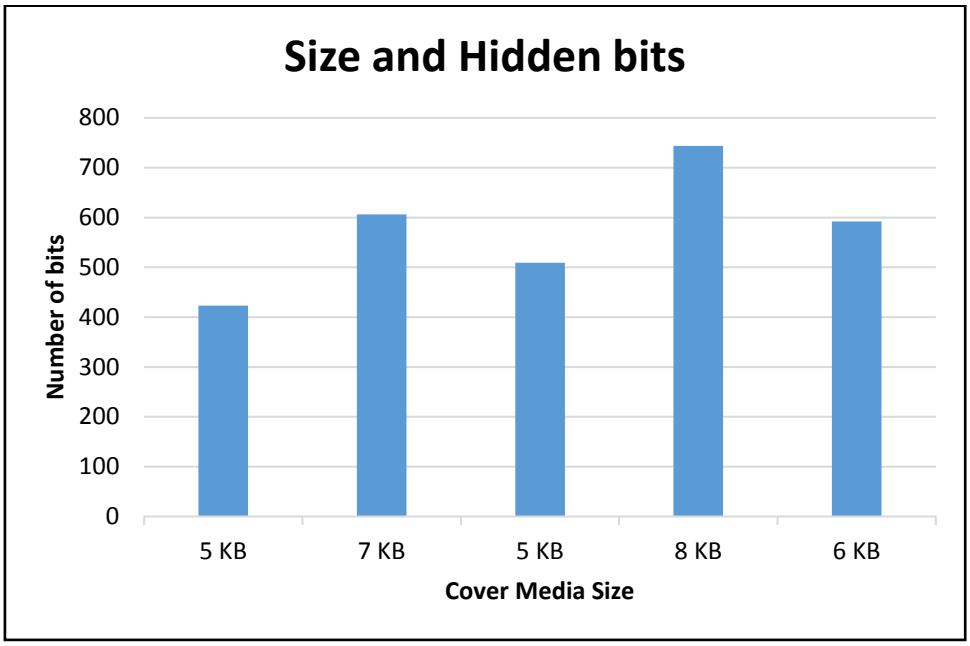

Figure 4.10: Relationship between Different Size and Amount of data can be hidden inside it.

Table 4.17 represents the file size and number of bits added to cover file. For hiding any message in a cover file, the space ratio must be greater than or equal to 1.

Table 4.17: Represents Different Cover files, Sizes, Number of Spaces, and Number of bits inserted at Second layer.

Cover File	Original Size	Number of Space	Secret Message	Number of Sentences Selected at First Layer	Number of bits hidden at Second Layer	Space Ratio
Cover1.txt	5 KB	423	A	2	2×8=16	26.4375
Cover2.txt	7 KB	606	AB	4	4×8=32	18.9375
Cover3.txt	5 KB	509	ABC	6	6×8=48	10.60417
Cover4.txt	8 KB	744	foot	8	8×8=64	11.625
Cover5.txt	6 KB	592	cover	10	10×8=80	7.4

In the above table, space ratio indicates that if cover files are suitable for hiding even more characters at first layer embedding, then still there is space enough to accommodate 25, 35, 28, 41 and 32 characters respectively at second layer embedding.

4.5 Robustness Analysis

This is a measure of the ability of the algorithm to retain the data embedded in the cover even after the cover text has been subjected to various changes. This analysis consists of three approaches. The first approach uses modification in line and paragraph spacing, the second approach involves re-typing same words of stego file, whereas the third approach is based on copying the whole contents of stego file to another empty file. The details of approaches are described below. Fig.4.11 shows the original stego text file generated which conceals the secret message **“cover”.**

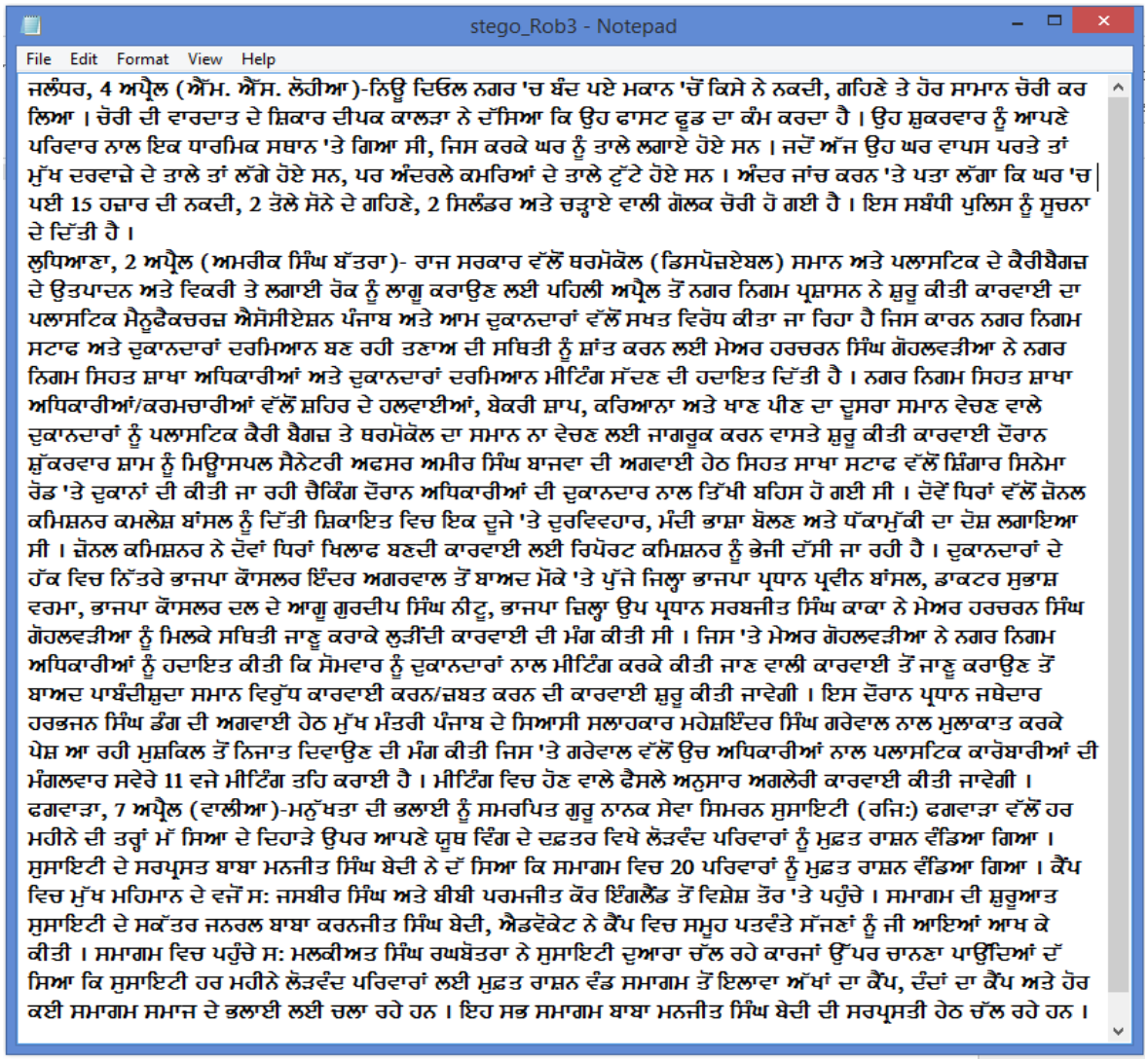

Figure 4.11: Original Stego-Text File

4.5.1 Modification in Line and Paragraph Spacing

In this approach of robustness test, line spacing between two lines and paragraph spacing before and after has been increased.

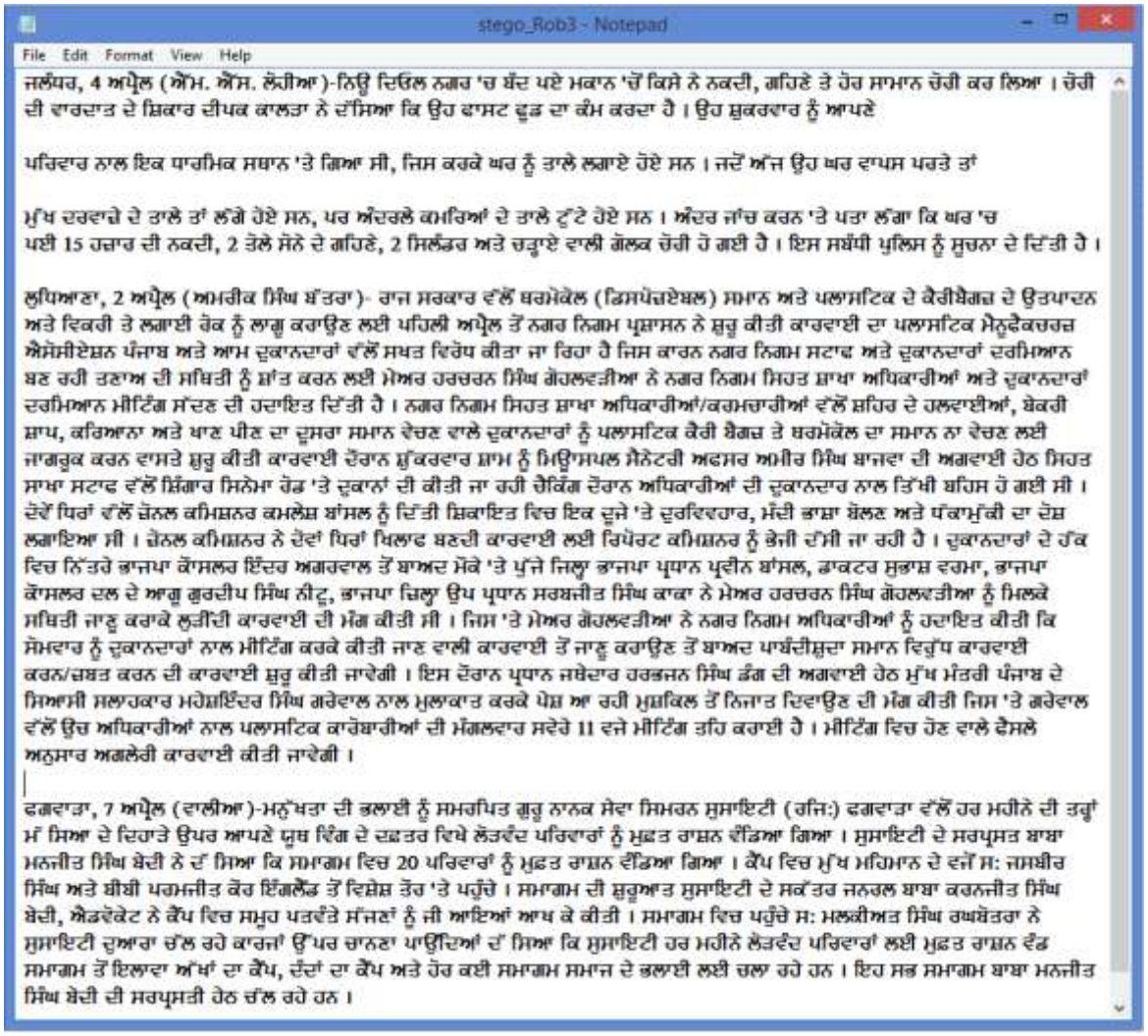

Figure 4.12: Modified Stego-Text File by Line and Paragraph Spacing

Since the contents of stego file do not change during formatting such as line and paragraph spacing. Hence, the data hidden in texts remains intact during these operations and hidden message **"cover"** has been extracted.

4.5.2 Retyping the Words in Stego File

In this approach, any characters or words are retyped without deleting the white space between words. It means that during the retyping process the white space should be untouched. In Fig. 4.13, the underlined words have been retyped. Since the feature or structure of the text in stego file does not change during retyping. Hence, the data hidden in texts remains intact during these operations.

4.5.3 Copying the Whole Contents of Original Stego File to Empty File

Since the contents of stego file do not change during copy and paste between files. Hence, the data hidden in texts remains intact during these operations and hidden message "cover" has been extracted.

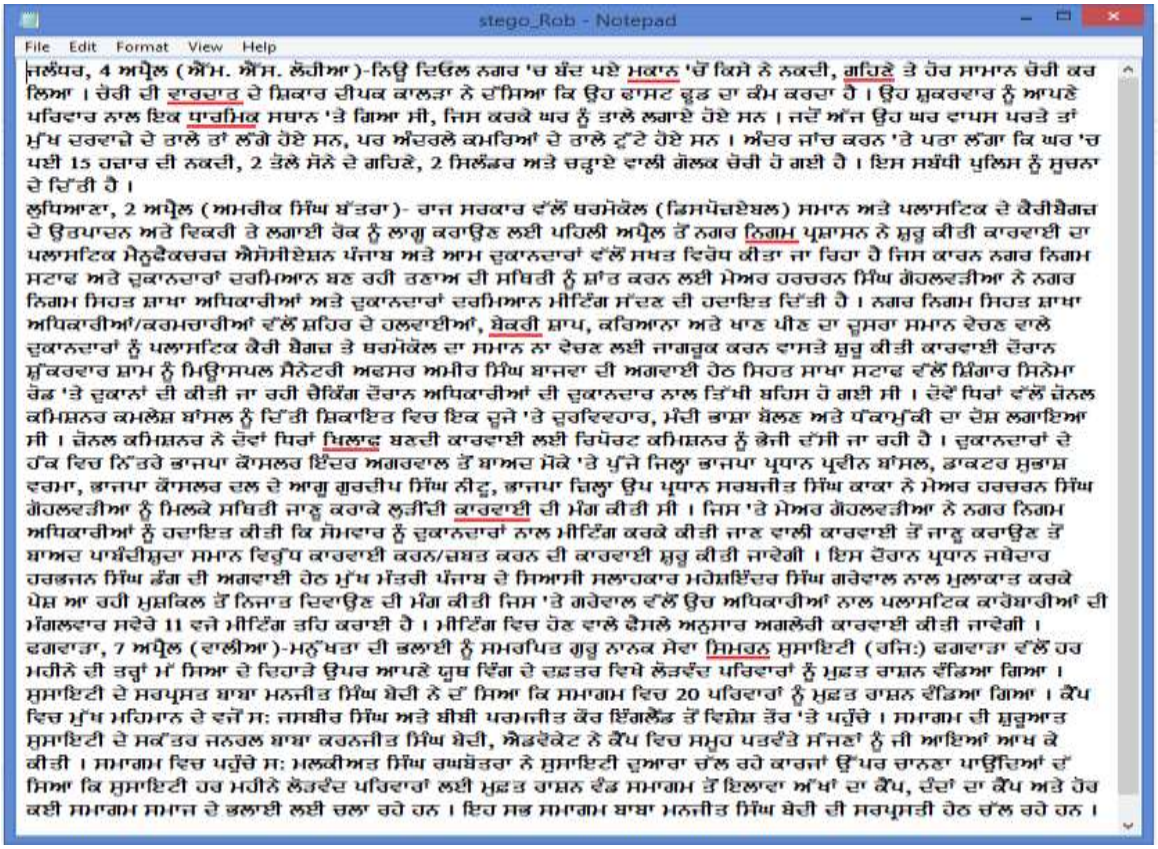

Figure 4.13: Stego-Text File in which Underlined Words have been Retyped

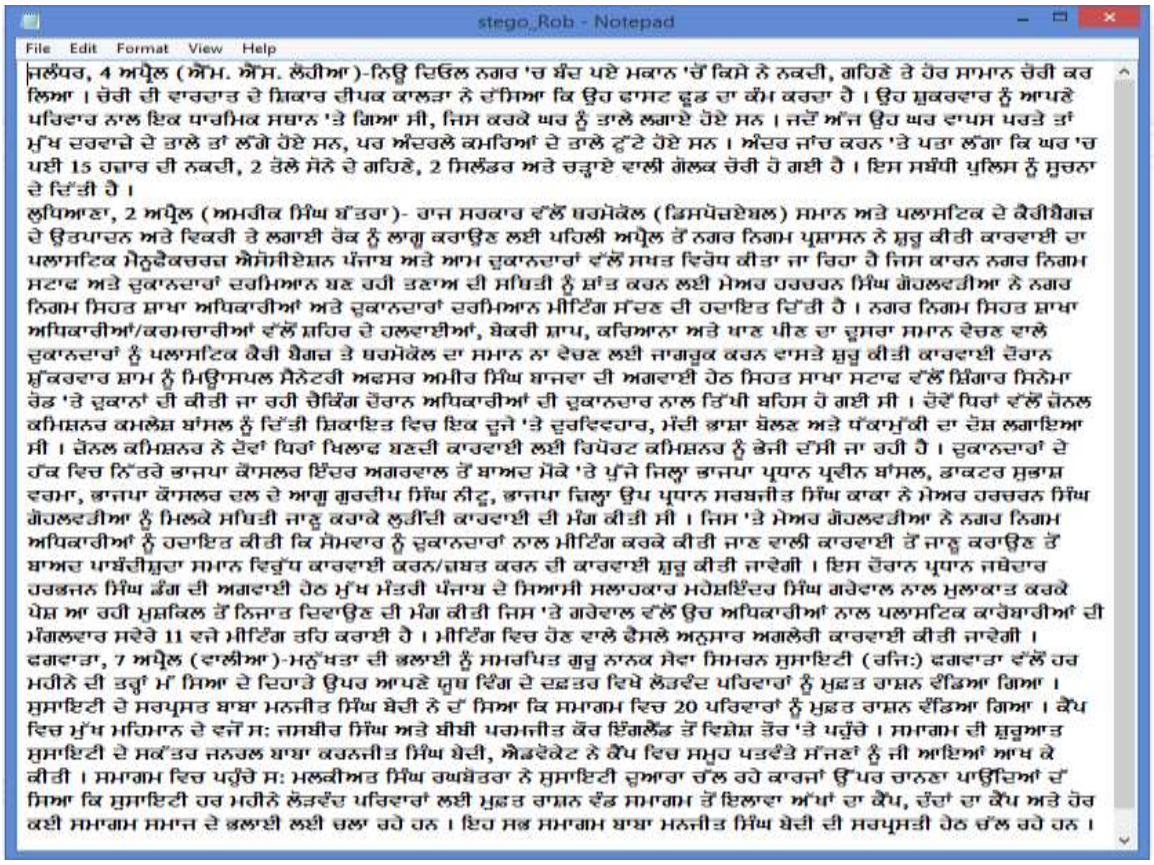

Figure 4.14: Contents Copied from Original Stego-Text File

CHAPTER 5

CONCLUSION

In this thesis, a new method to text steganography with an Indian local language, Punjabi has been developed. In the proposed approach, two layered embedding technique has been developed. The first layer embedding is done by considering the number of vowel, number of consonant, number of character and number of word in a sentence. In a sentence, 4 bits are concealed. For the embedding of any 4 bits possibilities in a sentence, some other characters along with the vowels and consonants are required. Thus, this technique is suitable for the Punjabi language as it has some characters in Gurumukhi character set that are neither vowel nor consonants. The same approach can be applicable for a language having the same specification as the Punjabi language. At second level, LRM and RLM character that are control characters and do not appear as glyphs on the screen are injected according to a binary string obtained from first layer embedding. An exactly reverse procedure is followed at the receiver side to retrieve the embedded message.

Since the cover media is pre-existing media, not system generated media, therefore a user will be free to use the cover media taken from any Punjabi newspaper, magazine or books and the syntax of a sentence and sequence of sentence both will be true grammatically. Even those opponent who have a very good command on the Punjabi language, will not be able to find that there is some information hidden in the stego text.

The Jaro score of comparing cover text and stego text is 0.995769, which means that they are almost identical. This property can be used to avoid stegoanalysis. It is found that the proposed method has very high hiding capacity without leaving any noticeable footprint for stegoanalysis. In addition, this method is robust to digital copy-paste operation, line and paragraph spacing modification and retyping the same characters or words. On the other side, there is some increase in the size of stego document that results from injecting the LMR and RLM character in space between two words. But the file format will not change which avoids the stegoanalysis.

However, once its applicability is known, it can easily be attacked. Hence, it is essential to keep the application of this approach secret, while using.

REFERENCES

Abduallah, W. M., Rahma, A. M. S., & Pathan, A. S. K. (2014). Mix column transform based on irreducible polynomial mathematics for color image steganography: A novel approach. *Computers & Electrical Engineering*, **40**: 1390-1404.

Aboalsamh, H. A., Mathkour, H., Mursi, M., & Assassa, G. M. (2008). Steganalysis of JPEG images: an improved approach for breaking the F5 algorithm. In *Proceedings of the 12th World Scientific and Engineering Academy and Society international conference on Computers,* 1011-1018, *Heraklion, Greece*.

Alla, K., & Prasad, R (2008). A Novel Hindi Text steganography using letter diacritics and its compound words. *International Journal of Computer Science and Network Security (IJCSNS*), **8**: 404-409.

Agarwal, M., (2013). Text Steganographic Approaches: A Comparison. *International Journal of Network Security & Its Applications,* **5**: 91-106

Bennett, K., (2004). Linguistic steganography: survey, analysis, and robustness concerns for hiding information in text. *CERIAS Tech. Report,* Purdue University, United States.

Bender, W., Gruhl, D., Morimoto, N., & Lu, A. (1996). Techniques for data hiding. *IBM systems journal,* **35:** 313-336.

Bassil, Y. (2012). A Generation-based Text Steganography Method using SQL Queries. *International Journal of Computer Applications,* **57**: 27-31.

Bhattacharyya, S., Banerjee, I., & Sanyal, G. (2010). A novel approach of secure text based steganography model using word mapping method (WMM). *International Journal of Computer and Information Engineering*, **4(2)**: 96-103.

Changder, S., Debnath, N. C., & Ghosh, D. (2009). A New approach to hindi text steganography by shifting matra. *International Conference on Advances in Recent Technologies in Communication and Computing, IEEE*, 199-202.

Changder, S., Ghosh, D., & Debnath, N. C. (2010). LCS based text steganography through Indian Languages. *3rd IEEE International Conference on Computer Science and Information Technology,* **8**: 53-57.

Codr, J. (2009). Unseen: An Overview of Steganography and Presentation of Associated Java Application C-Hide. *MS thesis,* Washington University, United States.

Dasare, A. J., & Dhore, M. L. (2015). Secured Approach for Hiding Data in MS Word Document Using MCDFF. *International Conference on Computing Communication Control and Automation, IEEE,* 296-300, Pune, India.

El-Abbadi, N. K., & El-Abbadi, A. N. Developing Measurements For Steganography Covers.<http://edu.uokufa.edu.iq/staff/dr.nidhal/researches/DEVELOPING%20ME

ASUREMENTS%20FOR%20STEGANOGRAPHY%20COVERS.doc>.Accessed 2015 July 30.

Garg, M. (2011). A novel text steganography technique based on html documents. *International Journal of Advanced Science and Technology*, ***35*:**129-138.

Inoue, S., Makino, K., Murase, I., Takizawa, O., Matsumoto, T., & Nakagawa, H. (2001). A proposal on information hiding methods using XML. *1st Workshop on NLP and XML,* 707-710, Tokyo.

Johnson, N. F., & Jajodia, S. (1998). Exploring steganography: Seeing the unseen. *IEEE Computer*, **31***:* 26-34.

Kabetta, H., & Dwiandiyanta, B. Y. (2011). Information Hiding in CSS: A Secure Scheme Text-Steganography using Public Key Cryptosystem. *International Journal on Cryptography and Information Security* (IJCSNS) **1**:13-22.

Krenn, R. (2004). Steganography and steganalysis. <http://www.krenn.nl/univ/cry/steg/article.pdf.>. Accessed 2015 July 16.

Kumar, A., & Pooja, K. (2010). Steganography-A data hiding technique. *International Journal of Computer Applications*, **9***:* 19-23.

Lehal, G. S., & Bhagat, M. (2004). Error pattern in Punjabi Typed Text. In *Proceedings of International Symposium on Machine Translation, NLP and TSS,* 128-141, New Delhi.

Low, S. H., Maxemchuk, N. F., Brassil, J. T., & O'Gorman, L. (1995). Document marking and identification using both line and word shifting. *Proceedings of 14th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '95*), 853-860, Boston, MA.

Morkel, T., Eloff, J. H., & Olivier, M. S. (2005). An overview of image steganography. *Proceeding of 5th Annual conference on Information Security*, 1-11, Sandton, South Africa.

Nosrati, M., Karimi, R., & Hariri, M. (2012). Audio Steganography: A Survey on Recent Approaches. *World Applied Programming***. *2****:*** 202-205.

Odeh, A., Elleithy, K., & Faezipour, M. (2013). Text Steganography Using Language Remarks. In *The American Society of Engineering Education*, Northfield, VT, USA.

Pandya, S. D. & Virparia, P. V. (2009). Testing various similarity metrics and their permutations with clustering approach in context free data cleaning. *Int. Journal of Computer Science and Security*, **3***:* 344-350.

Por, L. Y., & Delina, B. (2008). Information hiding: A new approach in text steganography. *7th WSEAS International Conference on Applied Computer & Applied Computational Science,* 689-695*, Hangzhou, China.*

Rahma, A. M. S., Bhaya, W. S., & Al-Nasrawi, D. A. (2013) Text Steganography Based On Unicode of Characters in Multilingual. *International Journal of Engineering Research and Applications,* **3**: 1153-1165

Shirali-Shahreza, M. (2008). Text steganography by changing words spelling. Proceeding of *10th International Conference on Advanced Communication Technology, IEEE*, **3**: 1912-1913.

Shirali-Shahreza, M. H., & Shirali-Shahreza, M. (2006). A new approach to Persian/Arabic text steganography. *Proceedings of 5th IEEE/ACIS International Conference on Computer and Information Science (ICIS-COMSAR'06*), 310-315, Honolulu, HI, USA.

Shirali-Shahreza, M. H., & Shirali-Shahreza, M. (2008). A new synonym text steganography. *International conference on intelligent information hiding and multimedia signal processing*, *IEEE*, 1524-1526, Harbin, China.

Singh, H., Singh, P. K., & Saroha, K. (2009). A survey on text based steganography. In Proceedings *of the 3rd National Conference; INDIACom-2009 Computing For Nation Development.* BVICAM, 26-27, New Delhi.

Uddin, M. P., Saha, M., Ferdousi, S. J., Ibn Afjal, M., & Marjan, M. A. (2014). Developing an efficient solution to information hiding through text steganography along with cryptography. In *9th International Forum on Strategic Technology, IEEE*, 14-17, Chittagong, Bangladesh.

Unicode consortium, (2015). *The Unicode Consortium.* Retrieved from unicode.org: <http://unicode.org/charts/PDF/U0A00.pdf >. Accessed 2015 Aug 10.

Veach,S., Williamson,S. (2015). Language Manuals for Culturally, Linguistically Diverse Communities. Home Page. <http://languagemanuals.weebly.com/uploads/4/8/5/3/4853169/punjabi.pdf > Accessed 2015 Aug 10.