Bengali Steganography using CALP with a Novel Bengali Word Processor

SOUVIK BHATTACHARYYA[1]
INDRADEEP BANERJEE[2]
GAUTAM SANYAL[3]

[1]Department of CSE, University Institute of Technology,
The University of Burdwan,West Bengal, India - 713104,
Email: souvik.bha@gmail.com

[2]Department of CSE, University Institute of Technology),
The University of Burdwan,West Bengal, India - 713104,
Email: ibanerjee2001@gmail.com

[3]Professor, Department of CSE and Dean (SW),
National Institute of Technology, Durgapur,
Mahatma Gandhi Avenue, West Bengal, India - 713209,
Email: nitgsanyal@gmail.com

Abstract. Recent years have witnessed the rapid development of the Internet and telecommunication techniques. Steganography is the art and science of communicating in a way which hides the existence of the communication. Considerable amount of work has been carried out by different researchers on steganography. In this work the authors propose a novel text steganography method for Bengali text generated through a new approach of Bengali text processor. Considering the structure of Bengali alphabet, secret message has been hidden through changing the pattern of Bengali alphabet letters. This approach uses the idea of structural and feature changing of the cover carrier which is visibly indistinguishable from the original to the human beings and may be modified for other India language also.

Keywords: Text Steganography, Unicode, Word Processor, Phonetic, Parser, Syllable, Bengali Glyphs, CALP (Changing in Alphabet Letter Patterns).

(Received July 19th, 2011 / Accepted December 26th, 2011)

1 Introduction

The technique of information hiding has been widely applied on various fields during the recent years [9] and the two major branches, viz. digital watermarking and steganography have been derived [15, 21]. Digital watermarking provides the protection of intellectual property, where as steganography concerns privacy of information under surveillance. As the goal of steganography is to hide the presence of a message and to create a covert channel, it can be seen as the complement of cryptography, whose goal is to hide the content of a message[2]. Steganography works have been carried out on different transmission media like images, video, text and audio [19] as shown in Figure 1.

Among them image steganography is the most popular of the lot. In this method the secret message is embedded into an image as noise to it [17, 26, 16]. In video steganography, same method may be used [34, 6, 7]. Audio steganography embeds the message into a cover audio file as noise at a frequency out of human hearing range [10]. Most difficult kind of steganography is text steganography or linguistic steganography because due to the lack of redundant information [19].

Figure 1: Types of Steganography

Reader may see [22, 33] for better understanding of the steganography methodology. Some Steganographic model with high security features has been presented in [3, 4, 5] and [28].

A block diagram of a generic text steganographic system is given in Figure 2.

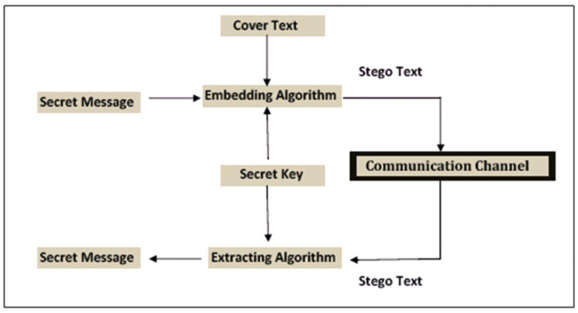


Figure 2: Generic form of Text Steganography

A message can be embedded in the cover text through an embedding algorithm to generate the stego text. During transmission of the stego text, it can be monitored by the unauthenticated viewers who will only notice the transmission of an innocuous-text without discovering the existence of the hidden message in it.

In this paper, a new method for Bengali text steganography is proposed. In this method cover text and secret message is generated through a novel bengali word processor which works through converting phonetic English word into Bengali word based on normal pronunciation. Stego text is generated by mapping the binary sequence of the secret message through texture/pattern changes of some alphabets of the cover text.

This paper is organized into the following sections. Section II describes some related works in text steganography. In Section III a new design approach of Bengali Word processor has been described. Section IV describes the text steganography method using CALP (Changing in Alphabet Letter Patterns). Section V describes the proposed Bengali Text Steganography methodology using CALP. Integer Wavelet Transform Technique describes in section VI. Algorithms of various functions are discussed in Section VII. Analysis of the Results are discussed in Section VIII. Computational Complexity of the algorithms of the system are described in Section IX. Mathematical formulation are described in section X. A comparative study of Bengali Steganography using CALP with some other existing methods are shown in section XI. The last section draws the conclusion.

2 Related Works

Text steganography can be broadly classified into three types - format-based, random and statistical generations and Linguistic method.

2.1 Format-based

Format-based methods use and change 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. A format-based text steganography method is open space method [22]. In this method extra white spaces are added into the text to hide information. A single space is interpreted as "0" and two consecutive spaces are interpreted as "1". Another two format-based methods are word shifting [7] and line shifting. Another method of hiding information in manipulation of white spaces between words and paragraph [25]. 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].

2.2 Random and statistical generation methods

Random and statistical generation methods are used to generate cover-text automatically according to the statistical properties of language. These methods uses example 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 [14]. The quality of the generated stego-message depends directly on the quality of the grammars used. Another approach to this type of method is to generate words having same statistical properties like word length and letter frequency of a word in the original message. The words generated are often without of any lexical value.

2.3 Linguistic method

The linguistic method [36] considers the linguistic properties of the text to modify it. The method uses linguistic structure of the message as a place to hide information. Syntactic method is a linguistic steganography method where some punctuation signs like comma (,) and full-stop (.) are placed in proper places in the document to embed a data.

2.4 Some Other Methods

Many researchers have suggested many methods for hiding information in text besides above three categories such as feature coding, text steganography by specific characters in words, abbreviations etc. [27] or by changing words spelling [20]. Some other methods like Text Steganography by Inter-word Spacing and Inter paragraph Spacing Approach [23] or Text Steganography by Using Letter Points and Extensions [11] or Text Steganography by Word Mapping Method(WMM) [29] or Text Steganography using Formatting Character Spacing [30] are also exist.

3 A Novel Bengali Word Processor

In this section the authors proposes an idea of a specific BENGALI WORD PROCESSOR [24] that search for the nearest word based on the users input from the database which makes the formation of a Bengali sentence faster and easier.This word processor has been designed by considering the following objectives:

- To help the user to form Bengali words from its normal pronunciation.
- To provide a system to understand natural human language using NLP.
- To help users by providing words closest to users input.
- To make sentences faster by searching and inserting words from Bengali word database.

The proposed Bengali word processor helps the user to make a word using normal pronunciation or better to say mapping Bengali words from normal English input according to pronunciation. The system consists of three parts: a GUI part, a Parser and a Database. In this proposed system when the user gives an input the given current input word is passed to the parser. The parser then parses this input word to its Bengali equivalent word or the subpart of a word accordingly. This parsed word or subpart of a word is passed to the database module and this module will search for the Bengali words very similar to the parsed word from the available database and it will provide the user with a list of probable words. From this list the user may be able to choose his/her desired word and will insert it to the current position or replace the word currently being edited by double-clicking on the word shown in the list-box. This process will continue whenever the user presses a key to edit a word. Figure 3 and figure 4 show the system frame work and GUI of the proposed Bengali Word Processor.

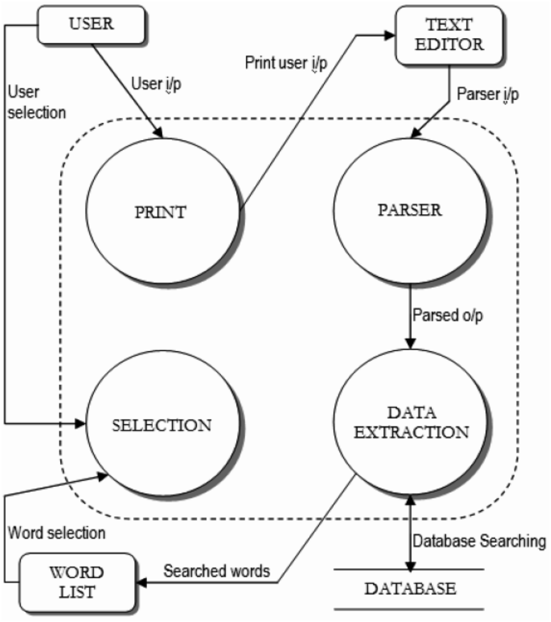
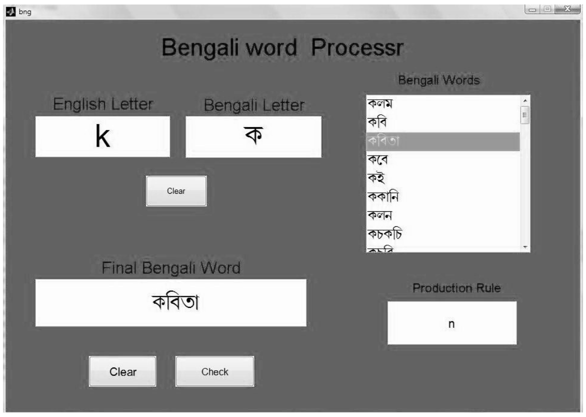


Figure 3: System Frame Work of the Word Processor



English Letter: k

Bengali Letter: ক

Bengali Words:
কলম
কবি
কবিতা
করে
কই
ককানি
কলন
কচকচি

Final Bengali Word: কবিতা

Production Rule: n

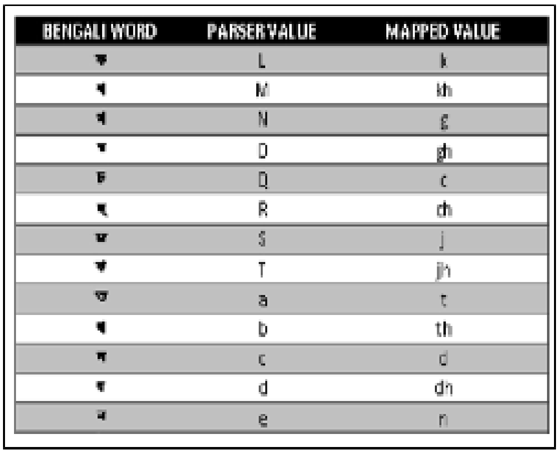
Figure 4: GUI of the System

3.1 User Interface

The interface system contains a text-box that will take the input from the user and for editing system will provide a list-box containing the predicted words and a data-grid having the available words from the database.

3.2 Parser

To convert a word from phonetic English to Bengali the concept of Natural Language Processing is used in the proposed system to form the Parser. A Parser is a program that divides the code into functional components. The input to this parser is taken through the keyboard in the form of phonetic English. Phonetic writing is a system that uses a unique symbol to represent each phone (sound) of the language or dialect. In case of phonetic Bengali the Bengali words is broken up into syllable (**unigrams**, **bigrams** etc) according to their pronunciations. In this word processor the Parser works as the main part of the application because it is the part which generates the ASCII character combinations corresponding to the Bengali word the users wished to write in actual. When the user inserts a word phonetically, the parser breaks the word into phonemes and replaces the phonemes with desired ASCII character combinations using some mappings stated in the parser. An example of the ASCII mappings used here is shown in the figure 5 below:



BENGALI WORD	PARSER VALUE	MAPPED VALUE
[illegible]	L	k
[illegible]	M	kh
[illegible]	N	g
[illegible]	D	gh
[illegible]	Q	c
[illegible]	R	ch
[illegible]	S	j
[illegible]	T	jh
[illegible]	a	t
[illegible]	b	th
[illegible]	c	d
[illegible]	d	dh
[illegible]	e	n

Figure 5: Table for mapping phonetic English to Bengali

The parser takes the word or the subpart of a word as its input from the user and parses it in a left to right manner and searches for syllables in the word in the same fashion. It searches for the syllables in a large to small manner that means it will first search for special sequence of characters, then bigrams and lastly for unigrams. This policy is taken to ensure that the best match is found to form the syllables more accurately from the word. Whenever the best match is found the syllable will be replaced with its corresponding Bengali ASCII character combinations. This process is repeated until the end of the word is reached. An example of word breaking has been shown in figure 6.

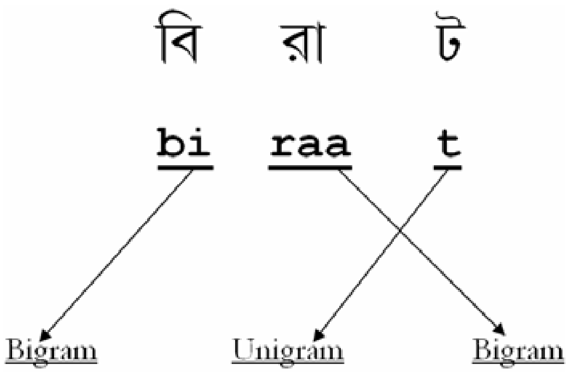


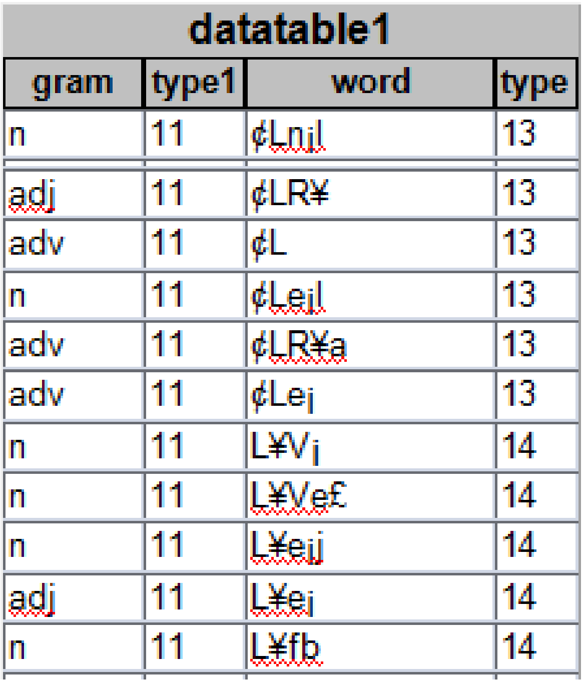
Figure 6: Example of breaking a word

After each time the parsing is complete the parsed Bengali word or the sequence of the ASCII characters are passed to the data extraction module.

3.3 Database Module

The Bengali word database contains a data table which contains two columns, one named word containing the Bengali words and another column named type left for future use which will show the type of words. The database stores the actual Bengali words in the database in their ASCII equivalent. To enter a word in the database we first have to know the exact ASCII values mapped to each of the Bengali glyphs to be used here. The structure of database module is shown in figure 7.

Using this certain knowledge, the Bengali words are converted to the sequence of ASCII characters and then this sequence is inserted to the database.While doing the opposite work, i.e. while extracting characters from the database, the parsed word is checked for those special characters and when found they are replaced accordingly. After the process of extraction is done, those data will then be copied to the Data-Grid. The Data-Grid is used here for easier access of the data. Finally the extracted data is inserted one-by-one in the List-Box for the user to access. In the List-Box the words are shown in Bengali and it allows the user to select the desired word and insert it to the current position by double-clicking on it. The word selected will replace the word currently being edited.



datatable1

gram	type1	word	type
n	11	¢Lnil	13
adj	11	¢LR¥	13
adv	11	¢L	13
n	11	¢Leil	13
adv	11	¢LR¥a	13
adv	11	¢Lej	13
n	11	L¥Vj	14
n	11	L¥Ve£	14
n	11	L¥eil	14
adj	11	L¥ej	14
n	11	L¥fb	14

Figure 7: Database Module

4 Text Steganography using Changing in Alphabet Letter Patter(CALP)

In this approach a new method for text steganography for English language is proposed [31, 32].In this method cover text and secret message is generated by the user. Stego text is formed by mapping the binary sequence of the secret message through texture/pattern changes of some alphabets of the cover text. Figures 8 and 9 below respectively show the mapping sequence for embedding 0s and 1s through the following pattern changes of the following alphabets of the cover text. These pattern changes have been incorporated using some unused symbols of the ASCII chart.

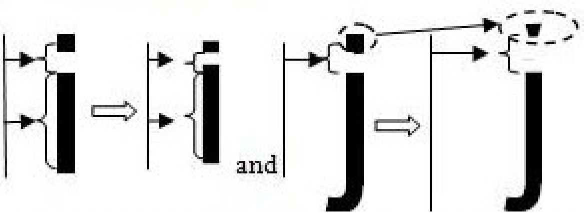


Figure 8: Mapping sequence for embedding ’0’ using CALP

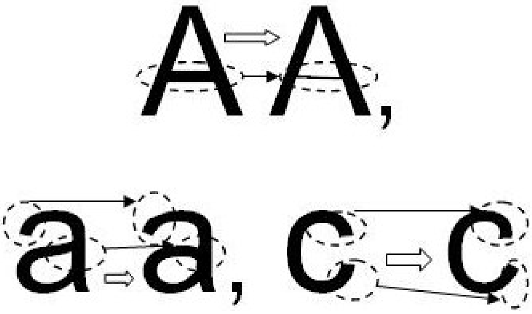


Figure 9: Mapping sequence for embedding ’1’ using CALP

5 Bengali Text Steganography using CALP

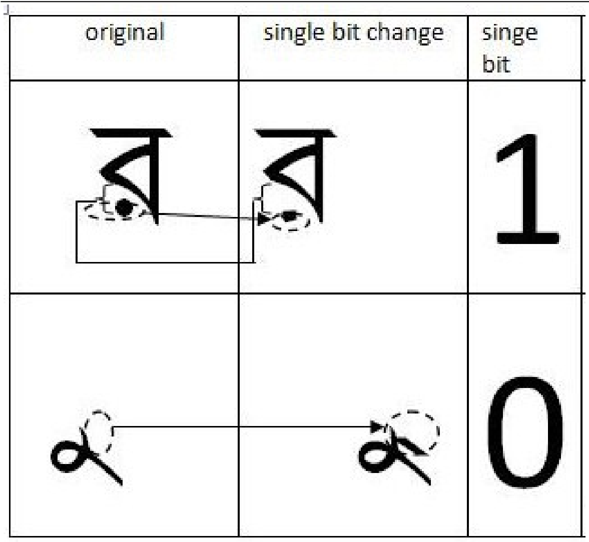
In this paper, a new method for text steganography for Bengali language is proposed. This method can be considered as the improved version of [31, 32] which has been used for Bengali text steganography method. In this method cover text and secret message is generated by the user. Stego text is formed by mapping the binary sequence of the secret message through texture/pattern changes of some alphabets of the cover text. There are two type of methodologies (i) Single-bit methodology (ii) Double-bit methodology.

Single-bit methodology: Here author deals with only 0 and 1 bits. Figure 10 below shows the mapping sequence for embedding 0s and 1s through the following pattern changes of the following alphabets of the cover text. These pattern changes have been incorporated using some unused symbols of the ASCII chart.

Double-bit methodology: Here the authors deal with four combinations of 0 and 1 bits. So in here author have four sequences of bits. They are ’00’, ’01’, ’10’ and ’11’. Figure 11 below shows the mapping sequence for embedding ’00’, ’01’, ’10’ and ’11’ and the difference of it with Single-bit methodology. Through the pattern changes of the following alphabets of the cover text. These pattern changes have been incorporated using some unused symbols of the ASCII chart.

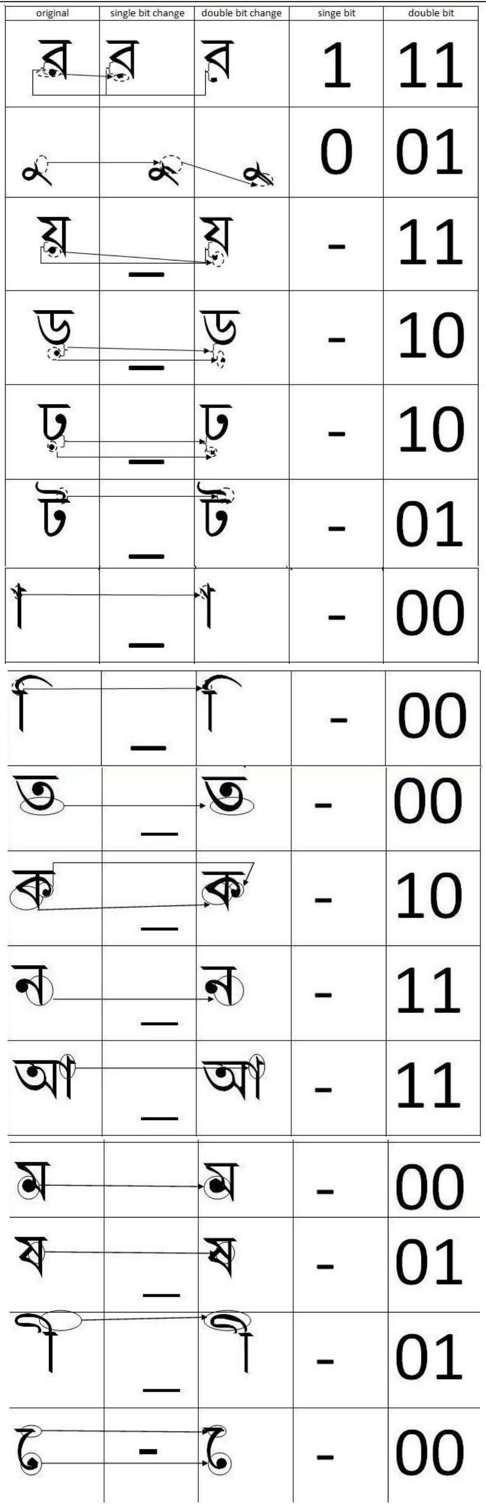
5.1 Solution Methodology

The **Sender Side** GUI consists of following two windows, one for the cover text generation and the other for the secret message generation. The user will be someone who is familiar with the process of information hiding and will have the knowledge of steganography systems and should be familiar with the proposed Bengali word processor. The user should be able to form a plain text as secret message, another text needs to be formed for use as the carrier (cover text). Before embedding the



original	single bit change	singe bit
র	র	1
ৎ	ৎ	0

Figure 10: Mapping sequence for embedding '0' and '1' for Single-bit method



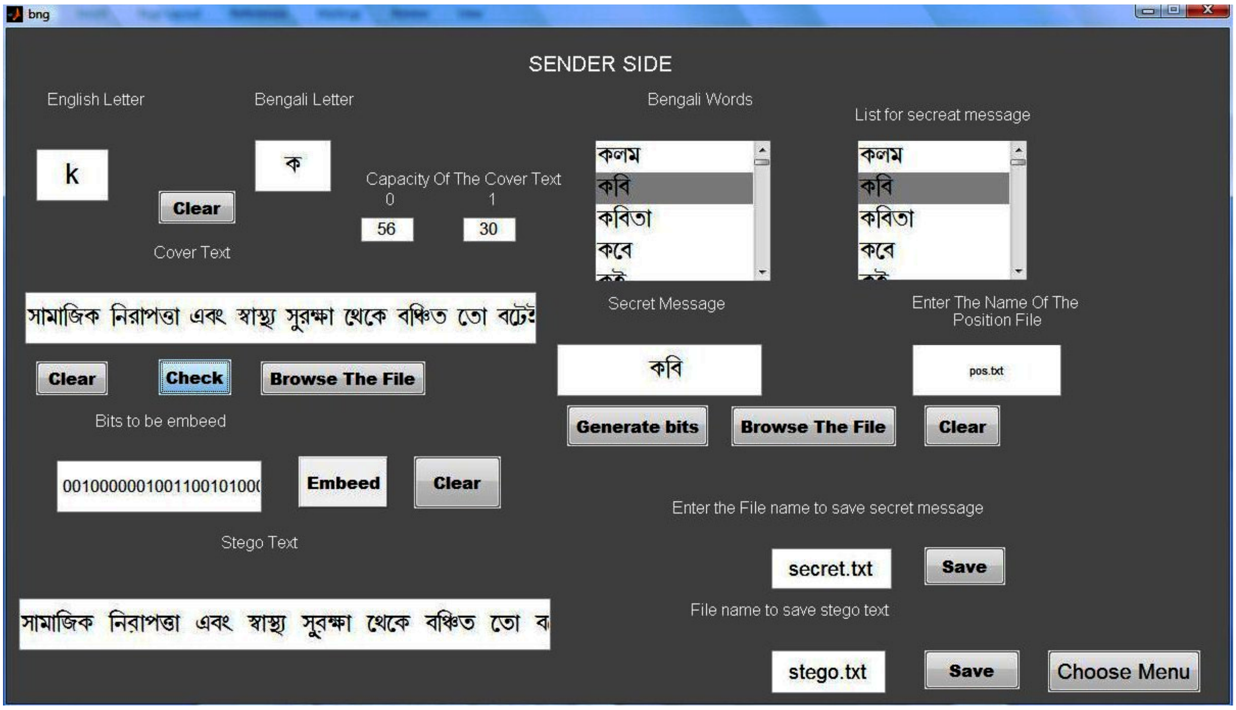
original	single bit change	double bit change	singe bit	double bit
র	র	র	1	11
ৎ	ৎ	ৎ	0	01
য়	_	য়	-	11
ড়	_	ড়	-	10
ঢ়	_	ঢ়	-	10
ট	_	ট	-	01
া	_	া	-	00
ি	_	ি	-	00
ত	_	ত	-	00
ক	_	ক	-	10
ন	_	ন	-	11
আ	_	আ	-	11
ম		ম	-	00
ষ	_	ষ	-	01
ী	_	ী	-	01
ে	-	ে	-	00

Figure 11: Mapping sequence for embedding '00', '01', '10' and '11' for Double-bit method

secret message will first be converted into unicode form which in turn encoded through integer wavelet transform. Finally the embedding method of the proposed system will be used to hide the encrypted version of secret message in the cover text to form the stego text. Secret Message will be extracted at the **Receiver Side** with the help of the different reverse order processes. Figures 12-15 (in the last pages of this article) show the different GUI for the proposed text steganography system for the Single-bit methodology and the Double-bit methodology respectively.

6 Integer Wavelet Transform

The lifting scheme is a technique for both designing wavelets and performing the discrete wavelet transform. Actually it is worthwhile to merge these steps and design the wavelet filters while performing the wavelet transform. The technique was introduced by Sweldens [25, 18]. The lifting scheme is an algorithm to calculate wavelet transforms in an efficient way. It is also a generic method to create so-called second-generation wavelets. They are much more flexible and can be used to define wavelet basis on an interval or on an irregular grid, or even on a sphere.The wavelet lifting scheme is a method for decomposing wavelet transform into a set of stages. An advantage of lifting scheme is that they do not require temporary storage in the calculation steps and have required less no of computation steps. The lifting procedure consists of three phases, namely, (i) split phase, (ii) predict phase and (iii) update phase



English Letter: k

Bengali Letter: ক

Capacity Of The Cover Text
0: 56
1: 30

Bengali Words: কলম, কবি, কবিতা, করে

List for secreat message: কলম, কবি, কবিতা, করে

Cover Text: সামাজিক নিরাপত্তা এবং স্বাস্থ্য সুরক্ষা থেকে বঞ্চিত তো বটেই

Secret Message: কবি

Enter The Name Of The Position File: pos.txt

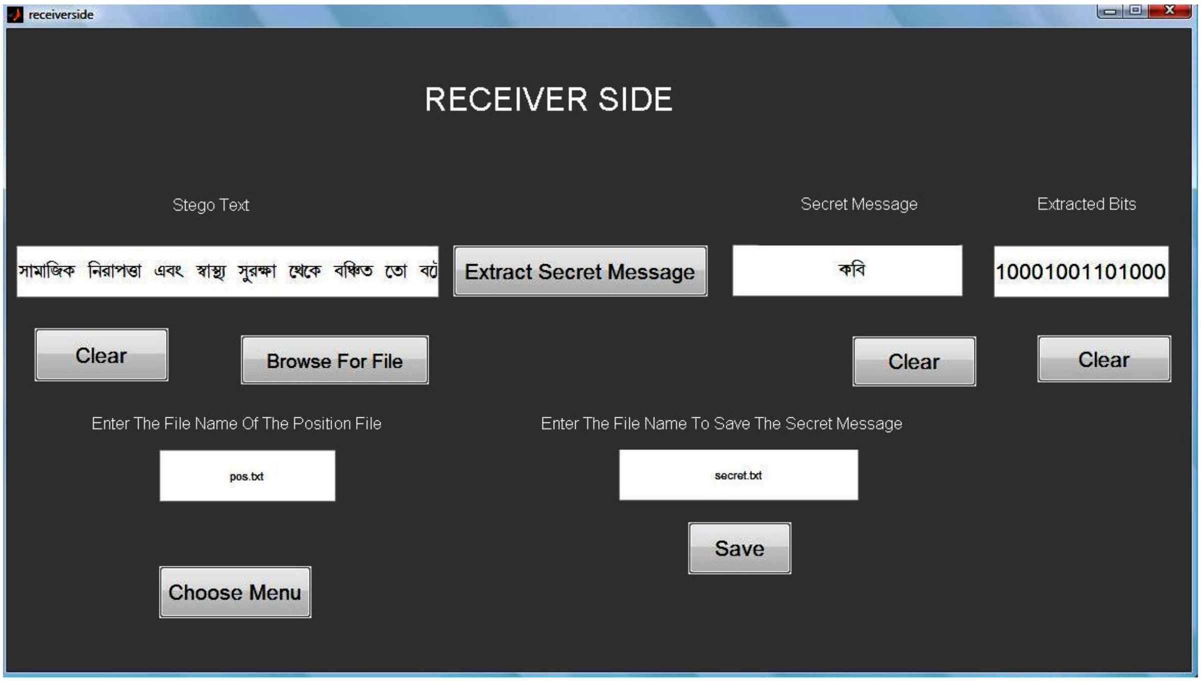
Bits to be embeed: 0010000001001100101000

Stego Text: সামাজিক নিরাপত্তা এবং স্বাস্থ্য সুবক্ষা থেকে বঞ্চিত তো ব

Enter the File name to save secret message: secret.txt

File name to save stego text: stego.txt

Figure 12: GUI of Sender-Side Single-bit methodology



Stego Text: সামাজিক নিরাপত্তা এবং স্বাস্থ্য সুরক্ষা থেকে বঞ্চিত তো ব

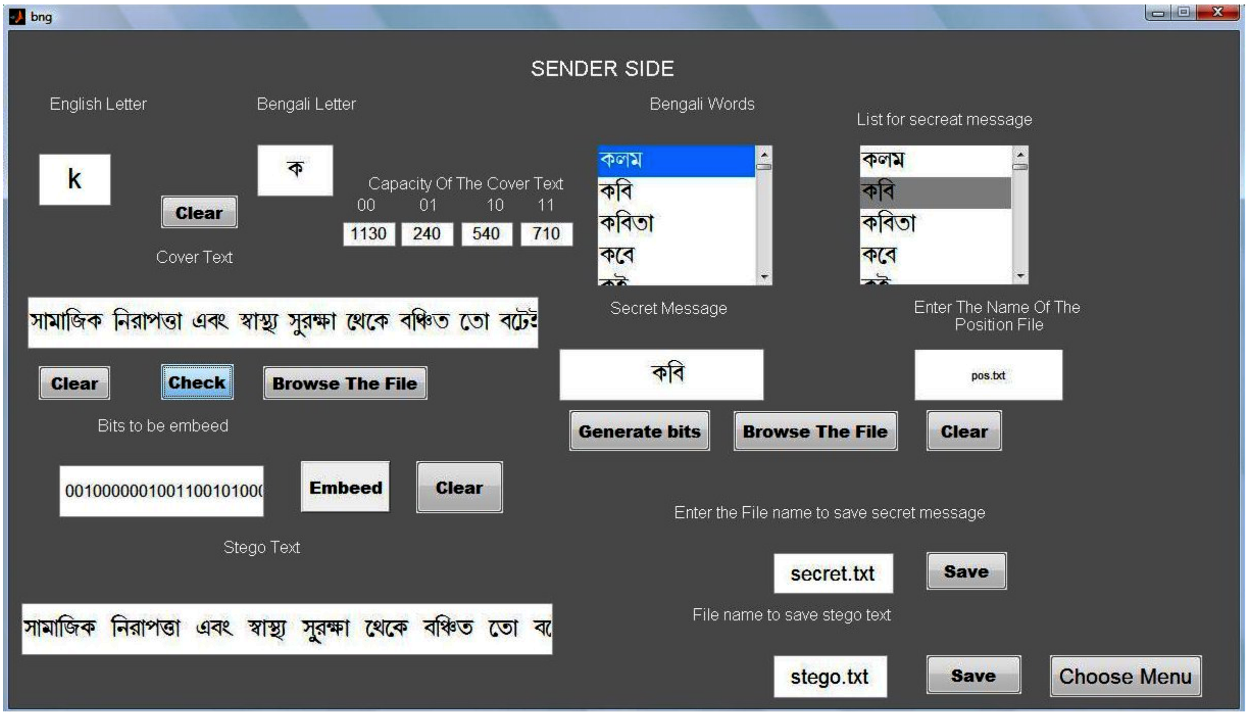
Secret Message: কবি

Extracted Bits: 10001001101000

Enter The File Name Of The Position File: pos.txt

Enter The File Name To Save The Secret Message: secret.txt

Figure 13: GUI of Receiver-Side Single-bit methodology



English Letter: k

Bengali Letter: ক

Capacity Of The Cover Text

00	01	10	11
1130	240	540	710

Bengali Words: কলম কবি কবিতা করে

List for secreat message: কলম কবি কবিতা করে

Cover Text: সামাজিক নিরাপত্তা এবং স্বাস্থ্য সুরক্ষা থেকে বঞ্চিত তো বটেই

Secret Message: কবি

Enter The Name Of The Position File: pos.txt

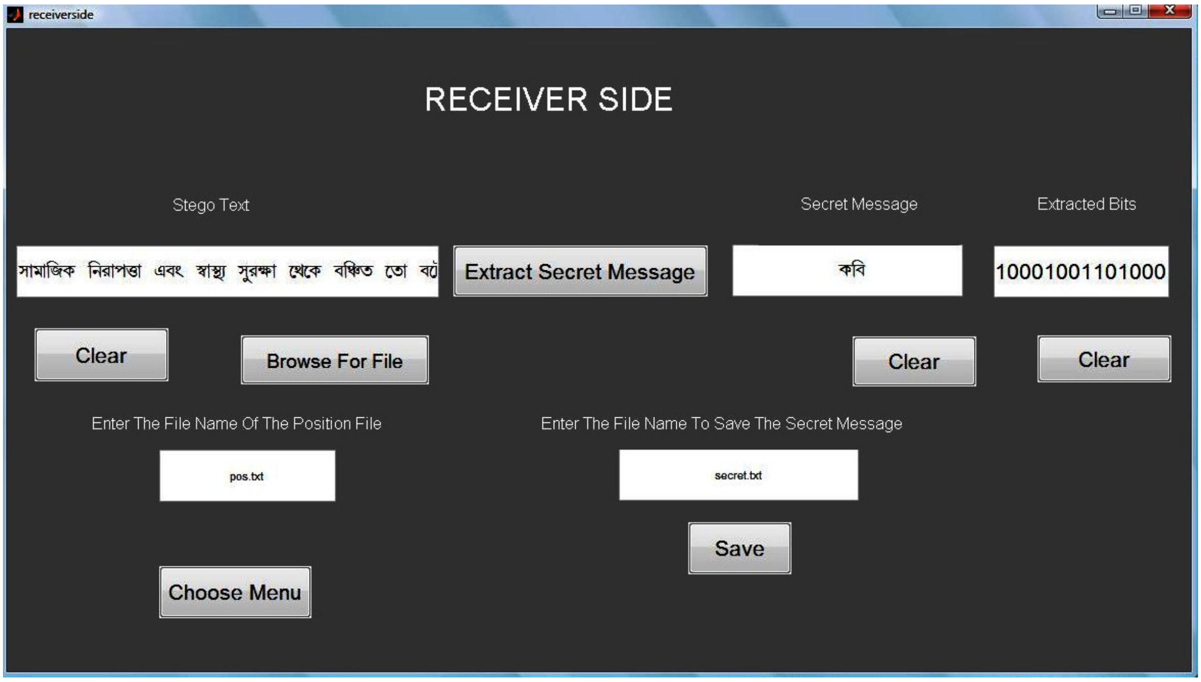
Bits to be embeed: 001000000100110010100(

Enter the File name to save secret message: secret.txt

Stego Text: সামাজিক নিরাপত্তা এবং স্বাস্থ্য সুরক্ষা থেকে বঞ্চিত তো ব

File name to save stego text: stego.txt

Figure 14: GUI of Sender-Side Double-bit methodology



Stego Text: সামাজিক নিরাপত্তা এবং স্বাস্থ্য সুরক্ষা থেকে বঞ্চিত তো ব

Secret Message: কবি

Extracted Bits: 10001001101000

Enter The File Name Of The Position File: pos.txt

Enter The File Name To Save The Secret Message: secret.txt

Figure 15: GUI of Receiver-Side Double-bit methodology

as shown in figure 16.

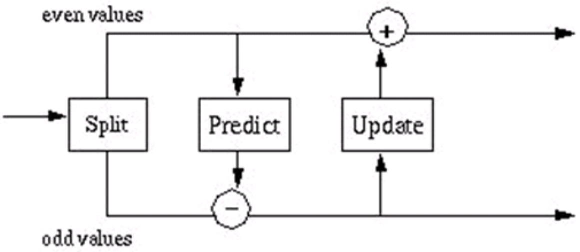


Figure 16: Lifting scheme forward wavelet transformation

Splitting: Split the signal x into even samples and odd samples: $x_{even} : s_i \leftarrow x_{2i}, x_{odd} : d_i \leftarrow x_{2i+1}$.

Prediction: Predict the odd samples using linear interpolation: $d_i \leftarrow d_i - \frac{(s_i + s_{i+1})}{2}$.

Update: Update the even samples to preserve the mean value of the samples: $s_i \leftarrow s_i + \frac{(d_{i-1} + d_i)}{4}$.

The output from the s channel provides a low pass filtered version of the input where as the output from the d channel provides the high pass filtered version of the input. The inverse transformed is obtained by reversing the order and the sign of the operations performed in the forward transform. Figure 17 shows the reverse process of the integer wavelet transform.

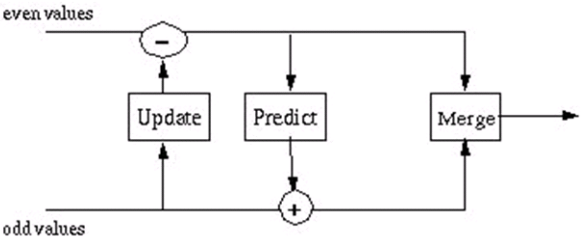


Figure 17: Lifting scheme inverse wavelet transformation

6.1 Lifting Scheme Haar Transform

In the lifting scheme version of the Haar transform, the prediction step predicts that the odd element will be equal to the even element. The difference between the predicted value (the even element) and the actual value of the odd element replaces the odd element. For the forward transform iteration j and element i, the new odd element, j+1,i would be: $odd_{j+1,i} = odd_{j,i} - even_{j,i}$. In the lifting scheme version of the Haar transform the update step replaces an even element with the average of the even / odd pair (e.g. the even element s_i and its odd successor s_{i+1}) is $even_{j+1,i} = \frac{(even_{j,i} + odd_{j,i})}{2}$. The original value of the $odd_{j,i}$ element has been replaced by the difference between this element and its even predecessor. The original value is $odd_{j,i} = even_{j,i} + odd_{j+1,i}$. Substituting this into the average $even_{j+1,i} = \frac{(even_{j,i} + even_{j,i} + odd_{j+1,i})}{2}$.

7 Algorithms

In this section various algorithm for embedding and extraction methodology for single bit and double bit has been discussed.

7.1 Algorithm for Message Encryption

- Generate the Secret Message.
- Convert the secret message into its corresponding UniCode form.
- Apply integer wavelet transformation technique in each 8 bit of the coded message to generate the approximate coefficient and detail coefficient.
- Convert the coefficients into binary form to generate the encoded form of the secret message.

7.2 Algorithm for Message Embedding using Single-Bit

Let COVER is cover text and STEGO is the string which consists of the stego text and MSG is the binary string of the secret encoded message and N is the no of elements in the MSG. Initially COVER and STEGO are the same. Set the counters i, j and r initialize to 1 and a, b, ca, cb are the counters which are initialized to zero. POS is an array which contains the positions of '0' bits of MSG and ENCRYPT is the function which encrypts a value.

- Select an appropriate COVER consisting of " ড" and " ব".
- Let k be the size of the COVER.Copy The contents of the COVER into STEGO.
- for j=1 to N
- if(($MSG(j) ==' 0'$) then $a = a + 1$
  else if(($MSG(j) ==' 1'$) then $b = b + 1$
- for i=1 to k
- if(($COVER(i) ==$ "ড") and ($cb < b - 1$))
  then $STEGO(i) =$ " ড় "
  else if(($COVER(i) ==$ "ব") and (($ca < a - 1$)))
  then $STEGO(i) =$ "র"

- if($cb < b - 1$)
  then STEGO(i)= " ৎ "
- for i=1 to N
- if ($MSG(i) == 0$)
  then POS(r)= ENCRYPT(i)
- increment r.
- END

7.3 Algorithm for Message Extraction in Single-bit methodology

Let STEGO is the stego text and MSG is the binary string of the secret message and N is the no. of elements in the STEGO and i and r be two arbitrary variables and j is initialize to 1. POS is an array which contains the positions of '0' bits of MSG and DEENCRYPT is the function which reverses the encrypt action on a value. LENGHT(MSG) gives the length of the secret message. SP is the number of element in the POS.

- for i=1 to N
- if($STEGO(i) ==$ " ৎ ") then $MSG(j) = 1$ and $j = j + 1$
  else if($STEGO(i) ==$ "র") then $MSG(j) = 0$ and $j = j + 1$
- for i =1 to LENGTH(MSG)
- $MSG(i) = 1$
- for i= 1 to SP
- $r = DEENCRYPT(MSG(SP))$
- $MSG(r) = 0$
- END

7.4 Algorithm for embedding in Double-bit methodology

Let COVER is cover text and STEGO is the string which consists of the stego text and MSG is the binary string of the secret message and N is the no of elements in the MSG. Initially COVER and STEGO are the same. Set two counters i, j and r initialize to 1 and a, b, c, d, ca, cb, cc, cd are the counters which are initialized to zero. POS is an array which contains the positions of '0' bits of MSG and ENCRYPT is the function which encrypts a value.

- Generate an appropriate COVER consisting of "ম" or "ত" or "ি" or " া " and "ী" or "ষ" or "ৎ " or "ট" and "ক" or "ে" or "ড়" or "ঢ়" and "ন" or "আ" or "র" or "য় ".
- Let k be the size of the COVER. Copy the contents of the COVER into STEGO.
- for j=1 to N
- if((MSG(j)== '0') and MSG(j+1)== '0') $a = a + 1$
  else if((MSG(j)== '0') and MSG(j+1)== '1') $b = b + 1$
  else if((MSG(j)== '1') and MSG(j+1)== '0') $c = c + 1$
  else if((MSG(j)== '1') and MSG(j+1)== '1') $d = d + 1$
- for i=1 to k
- if((COVER(i)== "ী" or "ষ" or "ৎ " or "ট") and ($cb < b - 1$)) then put STEGO(i)= " ী" or " ষ " or " ৎ " or " ট ")
  else if((COVER(i)== "ক" or "ে" or "ড়" or "ঢ়") and ($cc < c - 1$)) then put STEGO(i)= " ক " or " ে " or " ড় " or " ঢ় "
  else if((COVER(i)== " ন " or "আ" or " র " or "য় ") and ($cd < d - 1$))
  then put STEGO(i)=" ন " or " আ" or "র" or " য় "
  if((COVER(i)== "ম" or "ত" or "ি" or " া ") and ($ca < a - 1$)) then put STEGO(i)= " ম " or " ত " or " ি " or " া "
- for i=1 to N
- if ($MSG(i) == 0$) then POS(r) = ENCRYPT(i)
- End

7.5 Algorithm for Message Extraction using Double-Bit

Let STEGO is the stego text and MSG is the binary string of the secret message and N is the no. of elements in the STEGO and i and r be two arbitrary variables and j is initialize to 1. POS is an array which contains the positions of '0' bits of MSG and DEENCRYPT is the function which reverses the encrypt action on a value. LENGTH(MSG) gives the length of the secret message. SP is the number of element in the POS.

- for i=1 to N

- if(STEGO(i)== " ী" or " ষ " or " ঁ " or " ট ") then put $(MSG(j) = 0 and MSG(j+1) = 1)$ and $j = j + 2$

  else if(STEGO(i)== " ক " or " ে " or " ড় " or " ঢ় ") then put ($MSG(j) = 1 and MSG(j+1) = 0)$ and $j = j + 2$

  else if(STEGO(i)==" ন " or " আ" or "র" or " য় ") then put $(MSG(j) = 1 and MSG(j+1) = 1)$ and $j = j + 2$

  else if(STEGO(i)== " ম " or " ত " or " ি " or " া " then put $(MSG(j) = 0 and MSG(j + 1) = 0)$ and $j = j + 2$

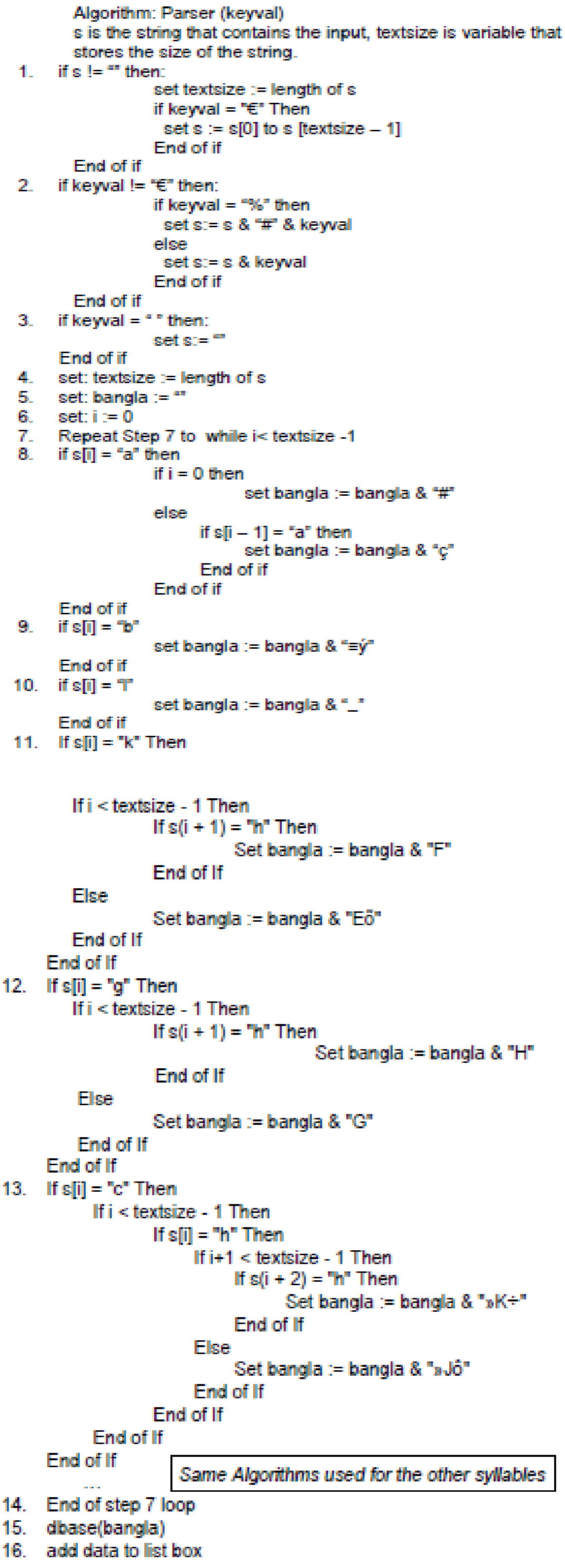
- for i =1 to LENGTH(MSG)
- MSG(i)=1
- for i= 1 to SP
- r= DEENCRYPT(MSG(SP))
- MSG(r)=0
- END

7.6 Algorithm for Bengali Word Processor

The main part of the proposed system is generating the parser. The basic concept of natural processing is used here to break the words into syllables. The syllables are recognized by using some predefined format to match with. The parser used here is a supervised system that contains the input-output patterns. The output corresponding to the input is searched from among the patterns available in the system. Figure 18 shows the algorithm of the Bengali Word Processor.

8 Analysis of the Results

There are mainly three aspects that should be taken into account while discussing the results of the proposed method of text steganography. They are security, capacity and robustness. The authors simulated the proposed system and the results are shown in the figures 19, 20, 21, 22, 23 and 24 respectively. This method satisfies both security aspects and hiding capacity requirements. It generates the stego text with minimum degradation which is not very revealing to people about the existence of any hidden data, maintaining its security to the eavesdroppers. Although the embedding capacity of the proposed method depends upon the cover text structure but the embedding capacity can be maximized by incorporating more no of alphabets through minor pattern

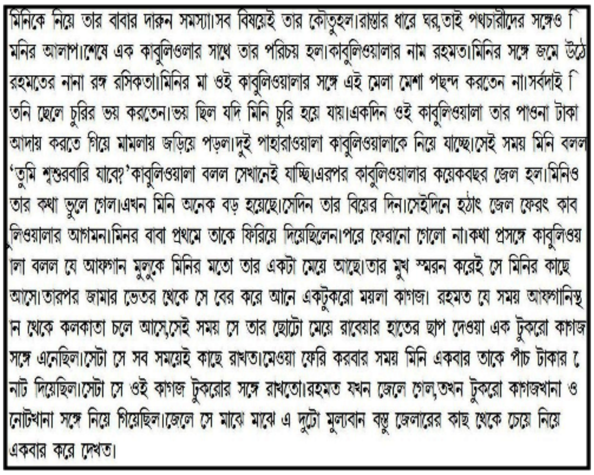


```
    Algorithm: Parser (keyval)
    s is the string that contains the input, textsize is variable that
    stores the size of the string.
1.  if s != "" then:
            set textsize := length of s
            if keyval = "€" Then
              set s := s[0] to s [textsize – 1]
            End of if
    End of if
2.  if keyval != "€" then:
            if keyval = "%" then
              set s:= s & "#" & keyval
            else
              set s:= s & keyval
            End of if
    End of if
3.  if keyval = " " then:
            set s:= ""
    End of if
4.  set: textsize := length of s
5.  set: bangla := ""
6.  set: i := 0
7.  Repeat Step 7 to  while i< textsize -1
8.  if s[i] = "a" then
            if i = 0 then
                    set bangla := bangla & "#"
            else
                if s[i – 1] = "a" then
                    set bangla := bangla & "ç"
                End of if
            End of if
    End of if
9.  if s[i] = "b"
            set bangla := bangla & "≡ý"
    End of if
10. if s[i] = "l"
            set bangla := bangla & "_"
    End of if
11. If s[i] = "k" Then

      If i < textsize - 1 Then
            If s(i + 1) = "h" Then
                    Set bangla := bangla & "F"
            End of If
      Else
            Set bangla := bangla & "Eõ"
      End of If
    End of If
12. If s[i] = "g" Then
      If i < textsize - 1 Then
            If s(i + 1) = "h" Then
                          Set bangla := bangla & "H"
            End of If
      Else
            Set bangla := bangla & "G"
      End of If
    End of If
13. If s[i] = "c" Then
        If i < textsize - 1 Then
            If s[i] = "h" Then
                If i+1 < textsize - 1 Then
                    If s(i + 2) = "h" Then
                        Set bangla := bangla & "»K÷"
                    End of If
                Else
                    Set bangla := bangla & "»Jô"
                End of If
            End of If
        End of If
    End of If          Same Algorithms used for the other syllables
        ...
14. End of step 7 loop
15. dbase(bangla)
16. add data to list box
```

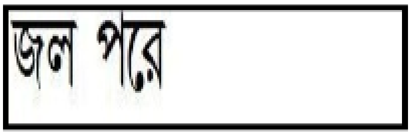
Figure 18: Algorithm for Bengali Word Processor

changes for mapping 0s and 1s for Single-bit methodology and ’00’, ’01’, ’10’ and ’11’ for double-bit methodology.



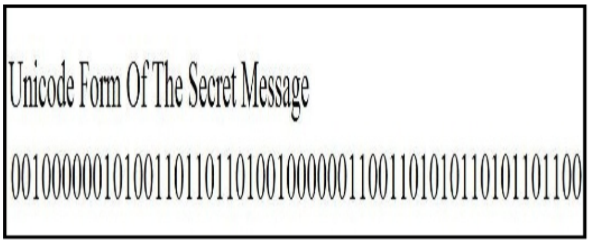
মিনিকে নিয়ে তার বাবার দারুন সমস্যা।সব বিষয়েই তার কৌতুহল।রাস্তার ধারে ঘর,তাই পথচারীদের সঙ্গেও ি
মনির আলাপ।শেষে এক কাবুলিওলার সাথে তার পরিচয় হল।কাবুলিওয়ালার নাম রহমত।মিনির সঙ্গে জমে উঠে
রহমতের নানা রঙ্গ রসিকতা।মিনির মা ওই কাবুলিওয়ালার সঙ্গে এই মেলা মেশা পছন্দ করতেন না।সর্বদাই ি
তনি ছেলে চুরির ভয় করতেন।ভয় ছিল যদি মিনি চুরি হয়ে যায়।একদিন ওই কাবুলিওয়ালা তার পাওনা টাকা
আদায় করতে গিয়ে মামলায় জড়িয়ে পড়ল।দুই পাহারাওয়ালা কাবুলিওয়ালাকে নিয়ে যাচ্ছে।সেই সময় মিনি বলল
‘তুমি শ্বশুরবারি যাবে?’কাবুলিওয়ালা বলল সেখানেই যাচ্ছি।এরপর কাবুলিওয়ালার কয়েকবছর জেল হল।মিনিও
তার কথা ভুলে গেল।এখন মিনি অনেক বড় হয়েছে।সেদিন তার বিয়ের দিন।সেইদিনে হঠাৎ জেল ফেরৎ কাব
লিওয়ালার আগমন।মিনর বাবা প্রথমে তাকে ফিরিয়ে দিয়েছিলেন।পরে ফেরানো গেলো না।কথা প্রসঙ্গে কাবুলিওয়
লা বলল যে আফগান মুলুকে মিনির মতো তার একটা মেয়ে আছে।তার মুখ স্মরন করেই সে মিনির কাছে
আসে।তারপর জামার ভেতর থেকে সে বের করে আনে একটুকরো ময়লা কাগজ। রহমত যে সময় আফগানিস্থ
ন থেকে কলকাতা চলে আসে,সেই সময় সে তার ছোটো মেয়ে রাবেয়ার হাতের ছাপ দেওয়া এক টুকরো কাগজ
সঙ্গে এনেছিল।সেটা সে সব সময়েই কাছে রাখত।মেওয়া ফেরি করবার সময় মিনি একবার তাকে পাঁচ টাকার ে
নাট দিয়েছিল।সেটা সে ওই কাগজ টুকরোর সঙ্গে রাখতো।রহমত যখন জেলে গেল,তখন টুকরো কাগজখানা ও
নোটখানা সঙ্গে নিয়ে গিয়েছিল।জেলে সে মাঝে মাঝে এ দুটো মুল্যবান বস্তু জেলারের কাছ থেকে চেয়ে নিয়ে
একবার করে দেখত।

Figure 19: Cover Text for both Single-bit and Double-bit methodologies



জল পরে

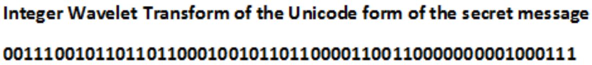
Figure 20: Secret Message for both Single-bit and Double-bit methodologies



Unicode Form Of The Secret Message

0010000001010011011011010010000001100110101011010101101100

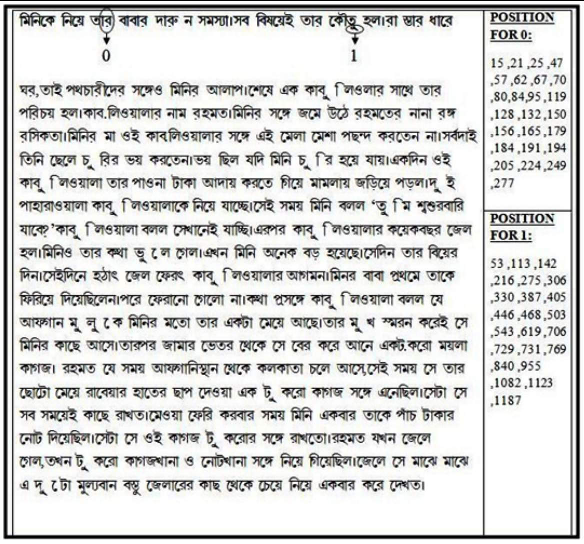
Figure 21: Unicode form of The Secret Message



Integer Wavelet Transform of the Unicode form of the secret message

00111001011011011000100101101100001100110000000000001000111

Figure 22: Encrypted form of The Secret Message



মিনিকে নিয়ে তার বাবার দারু ন সমস্যা।সব বিষয়েই তার কৌতু হল।রা স্তার ধারে

ঘর,তাই পথচারীদের সঙ্গেও মিনির আলাপ।শেষে এক কাবু লিওলার সাথে তার
পরিচয় হল।কাব.লিওয়ালার নাম রহমত।মিনির সঙ্গে জমে উঠে রহমতের নানা রঙ্গ
রসিকতা।মিনির মা ওই কাবলিওয়ালার সঙ্গে এই মেলা মেশা পছন্দ করতেন না।সর্বদাই
তিনি ছেলে চু রির ভয় করতেন।ভয় ছিল যদি মিনি চু রি হয়ে যায়।একদিন ওই
কাবু লিওয়ালা তার পাওনা টাকা আদায় করতে গিয়ে মামলায় জড়িয়ে পড়ল।দু ই
পাহারাওয়ালা কাবু লিওয়ালাকে নিয়ে যাচ্ছে।সেই সময় মিনি বলল ‘তু মি শ্বশুরবারি
যাবে?’কাবু লিওয়ালা বলল সেখানেই যাচ্ছি।এরপর কাবু লিওয়ালার কয়েকবছর জেল
হল।মিনিও তার কথা ভু লে গেল।এখন মিনি অনেক বড় হয়েছে।সেদিন তার বিয়ের
দিন।সেইদিনে হঠাৎ জেল ফেরৎ কাবু লিওয়ালার আগমন।মিনর বাবা প্রথমে তাকে
ফিরিয়ে দিয়েছিলেন।পরে ফেরানো গেলো না।কথা প্রসঙ্গে কাবু লিওয়ালা বলল যে
আফগান মু লু কে মিনির মতো তার একটা মেয়ে আছে।তার মু খ স্মরন করেই সে
মিনির কাছে আসে।তারপর জামার ভেতর থেকে সে বের করে আনে একট.করো ময়লা
কাগজ। রহমত যে সময় আফগানিস্থান থেকে কলকাতা চলে আসে,সেই সময় সে তার
ছোটো মেয়ে রাবেয়ার হাতের ছাপ দেওয়া এক টু করো কাগজ সঙ্গে এনেছিল।সেটা সে
সব সময়েই কাছে রাখত।মেওয়া ফেরি করবার সময় মিনি একবার তাকে পাঁচ টাকার
নোট দিয়েছিল।সেটা সে ওই কাগজ টু করোর সঙ্গে রাখতো।রহমত যখন জেলে
গেল,তখন টু করো কাগজখানা ও নোটখানা সঙ্গে নিয়ে গিয়েছিল।জেলে সে মাঝে মাঝে
এ দু টো মুল্যবান বস্তু জেলারের কাছ থেকে চেয়ে নিয়ে একবার করে দেখত।

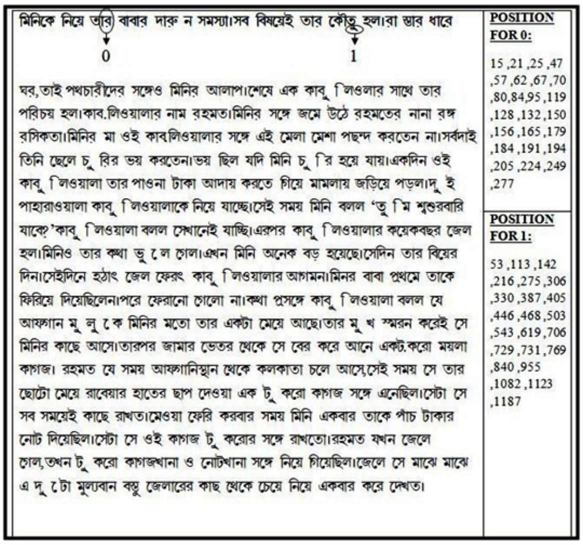
POSITION FOR 0:

15 ,21 ,25 ,47
,57 ,62 ,67 ,70
,80,84,95 ,119
,128 ,132 ,150
,156 ,165 ,179
,184 ,191 ,194
,205 ,224 ,249
,277

POSITION FOR 1:

53 ,113 ,142
,216 ,275 ,306
,330 ,387 ,405
,446 ,468 ,503
,543 ,619 ,706
,729 ,731 ,769
,840 ,955
,1082 ,1123
,1187

Figure 23: Stego Text for Single-bit methodology with the embedding positions



মিনিকে নিয়ে তার বাবার দারু ন সমস্যা।সব বিষয়েই তার কৌতু হল।রা স্তার ধারে

ঘর,তাই পথচারীদের সঙ্গেও মিনির আলাপ।শেষে এক কাবু লিওলার সাথে তার
পরিচয় হল।কাব.লিওয়ালার নাম রহমত।মিনির সঙ্গে জমে উঠে রহমতের নানা রঙ্গ
রসিকতা।মিনির মা ওই কাবলিওয়ালার সঙ্গে এই মেলা মেশা পছন্দ করতেন না।সর্বদাই
তিনি ছেলে চু রির ভয় করতেন।ভয় ছিল যদি মিনি চু রি হয়ে যায়।একদিন ওই
কাবু লিওয়ালা তার পাওনা টাকা আদায় করতে গিয়ে মামলায় জড়িয়ে পড়ল।দু ই
পাহারাওয়ালা কাবু লিওয়ালাকে নিয়ে যাচ্ছে।সেই সময় মিনি বলল ‘তু মি শ্বশুরবারি
যাবে?’কাবু লিওয়ালা বলল সেখানেই যাচ্ছি।এরপর কাবু লিওয়ালার কয়েকবছর জেল
হল।মিনিও তার কথা ভু লে গেল।এখন মিনি অনেক বড় হয়েছে।সেদিন তার বিয়ের
দিন।সেইদিনে হঠাৎ জেল ফেরৎ কাবু লিওয়ালার আগমন।মিনর বাবা প্রথমে তাকে
ফিরিয়ে দিয়েছিলেন।পরে ফেরানো গেলো না।কথা প্রসঙ্গে কাবু লিওয়ালা বলল যে
আফগান মু লু কে মিনির মতো তার একটা মেয়ে আছে।তার মু খ স্মরন করেই সে
মিনির কাছে আসে।তারপর জামার ভেতর থেকে সে বের করে আনে একট.করো ময়লা
কাগজ। রহমত যে সময় আফগানিস্থান থেকে কলকাতা চলে আসে,সেই সময় সে তার
ছোটো মেয়ে রাবেয়ার হাতের ছাপ দেওয়া এক টু করো কাগজ সঙ্গে এনেছিল।সেটা সে
সব সময়েই কাছে রাখত।মেওয়া ফেরি করবার সময় মিনি একবার তাকে পাঁচ টাকার
নোট দিয়েছিল।সেটা সে ওই কাগজ টু করোর সঙ্গে রাখতো।রহমত যখন জেলে
গেল,তখন টু করো কাগজখানা ও নোটখানা সঙ্গে নিয়ে গিয়েছিল।জেলে সে মাঝে মাঝে
এ দু টো মুল্যবান বস্তু জেলারের কাছ থেকে চেয়ে নিয়ে একবার করে দেখত।

POSITION FOR 0:

15 ,21 ,25 ,47
,57 ,62 ,67 ,70
,80,84,95 ,119
,128 ,132 ,150
,156 ,165 ,179
,184 ,191 ,194
,205 ,224 ,249
,277

POSITION FOR 1:

53 ,113 ,142
,216 ,275 ,306
,330 ,387 ,405
,446 ,468 ,503
,543 ,619 ,706
,729 ,731 ,769
,840 ,955
,1082 ,1123
,1187

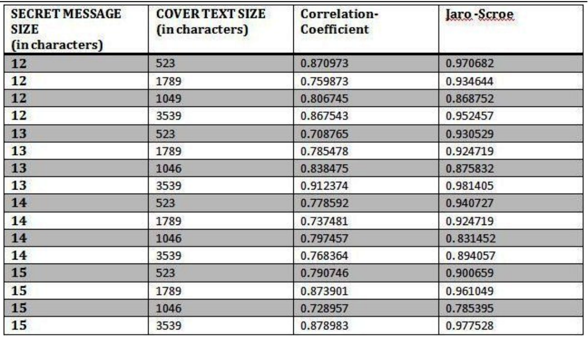
Figure 24: Stego Text for Double-bit methodology with the embedding positions

8.1 Similarity Measure Of The Cover Text And Stego Text Through Correlation

The most familiar measure of dependence between two quantities is the Pearson product-moment correlation coefficient [8], or "Pearson’s correlation." It is obtained by dividing the covariance of the two variables by the product of their standard deviations. Karl Pearson developed the coefficient from a similar but slightly different idea by Francis Galton. The Pearson correlation is +1 in the case of a perfect positive (increasing) linear relationship (correlation), -1 in the case of a perfect decreasing (negative) linear relationship (anti correlation), and some value between -1 and 1 in all other cases, indicating the degree of linear dependence between the variables. As it approaches zero there is less of a relationship (closer to uncorrelated). The closer the coefficient is to either -1 or 1, the stronger the correlation between the variables. If the variables are independent, Pearson's correlation coefficient is 0, but the converse is not true because the correlation coefficient detects only linear dependencies between two variables. If we have a series of n measurements of X and Y written as x_i and y_i, where i = 1, 2, 3, ..., n, then the sample correlation coefficient can be used in Pearson correlation r between X and Y. The sample correlation coefficient is written as $r_{xy} = \left(\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})\right) / ((n-1)s_x s_y)$, where $\overline{x}$ and $\overline{y}$ are the sample means of X and Y, s_x and s_y are the sample standard deviations of X and Y.

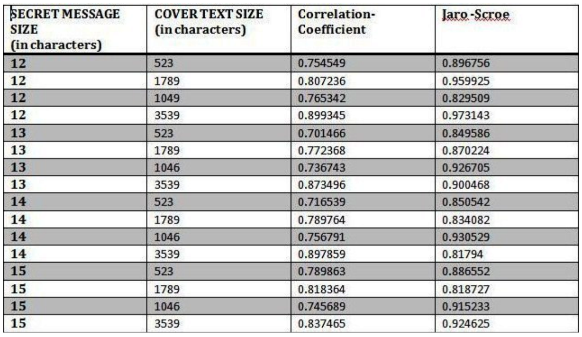
8.2 Similarity Measure Of The Cover Text And Stego Text Through Jaro Winkler Distance

For comparing the similarity between cover text and the stego text, the Jaro-Winkler distance for measuring similarity between two strings has been computed. The Jaro-Winkler distance [35] is a measure of similarity between two strings. It is a variant of the Jaro distance metric [12, 13] and mainly used in the area of record linkage (duplicate detection). The higher the Jaro-Winkler distance for two strings is, the more similar the strings are. The score is normalized such that 0 equates to no similarity and 1 is an exact match. The Jaro distance metric states that given two strings s_1 and s_2 their distance d_j is $d_j = \frac{1}{3}\left[\frac{m}{|s_1|} + \frac{m}{|s_2|} + \frac{m-t}{m}\right]$,where m is the number of matching characters and t is the number of transpositions.Two characters from s_1 and s_2 respectively are considered matching only if they are not farther than $\left\lfloor \frac{\max[|s_1|,|s_2|]}{2} \right\rfloor - 1$. Each character of s_1 is compared with all its matching characters in s_2 . The number of matching (but different sequence order) characters divided by two defines the number of transpositions. Figures 25 and 26 shows the Correlation coefficient and Jaro score for various size of cover text along with various size of the secret message of the proposed Single-bit methodology and Double-bit methodology.



SECRET MESSAGE SIZE (in characters)	COVER TEXT SIZE (in characters)	Correlation-Coefficient	Jaro -Scroe
12	523	0.870973	0.970682
12	1789	0.759873	0.934644
12	1049	0.806745	0.868752
12	3539	0.867543	0.952457
13	523	0.708765	0.930529
13	1789	0.785478	0.924719
13	1046	0.838475	0.875832
13	3539	0.912374	0.981405
14	523	0.778592	0.940727
14	1789	0.737481	0.924719
14	1046	0.797457	0.831452
14	3539	0.768364	0.894057
15	523	0.790746	0.900659
15	1789	0.873901	0.961049
15	1046	0.728957	0.785395
15	3539	0.878983	0.977528

Figure 25: Single-Bit Methodology Parameters



SECRET MESSAGE SIZE (in characters)	COVER TEXT SIZE (in characters)	Correlation-Coefficient	Jaro -Scroe
12	523	0.754549	0.896756
12	1789	0.807236	0.959925
12	1049	0.765342	0.829509
12	3539	0.899345	0.973143
13	523	0.701466	0.849586
13	1789	0.772368	0.870224
13	1046	0.736743	0.926705
13	3539	0.873496	0.900468
14	523	0.716539	0.850542
14	1789	0.789764	0.834082
14	1046	0.756791	0.930529
14	3539	0.897859	0.81794
15	523	0.789863	0.886552
15	1789	0.818364	0.818727
15	1046	0.745689	0.915233
15	3539	0.837465	0.924625

Figure 26: Double-bit methodology parameters

9 Computational Complexity Analysis

In this section computational complexity in terms of time and space for various algorithm of embedding and extraction methodology for single bit and double bit has been discussed.

9.1 Time Complexity At Sender Side for Single Bit /Double Bit Method

- Calculation the no of 1 bit sequence (i.e. 0 and 1) or 2 bit sequence (i.e. 00, 01, 10 and 11) in the **secret message**. Here the computation time is = **O(n)**[n is the total no of bits in the **secret message**].

- Computing the no of letters in the **cover text** individually that are used for embedding the 1 bit sequence or 2 bit sequence. Computation time is **=O(n)**[n is the total no of bits in the**cover text**].

- Recording or storing positions of '0' bits of the secret message into **array(pos)** and to encrypt those values. Computation time for storing '0' bits in the **pos array is=O(n).** Computation time for performing **ENCRYPT Function is=O(1).**

- Embedding the **secret message** into cover text. Computation time for embedding **secret message is=O(n).**

Total time complexity for **embedding secret message in cover text is O(n).**

9.2 Time Complexity At Receiver Side for Single Bit/Double Bit Method

- Retrieving the encrypted positions for storing ’0’ bit and to decrypt those values. Computation time for retrieving the **encrypted bits=O(n).** Computation time for performing **DECRYPT function is=O(1).**
- Extracting the embedded **secret message** bits from the **stego text.** Computation time for extracting **secret message is=O(n)**.

Total time complexity for **extracting secret message from stego text is O(n).**

9.3 Space Complexity at Sender Side for Single Bit/ Double Bit Method

- Storing the cover text in a string **COVER** whose length is n. Space complexity **O(n)**
- Storing the **secret message** in 1 or 2 bit format. Space complexity **O(n)**
- Embedding the **secret message** into cover text **COVER**.
  Space complexity **O(n)**
- Recording or storing positions of ’0’ bits of the secret message into **array(pos)** and to encrypt those values. Space complexity **O(n)**

Total space complexity for **Embedding procedure is O(n).**

9.4 Space Complexity At Receiver Side for Single Bit/Double Bit Method

- Retrieving the encrypted positions for storing ’0’ bit and to decrypt those values. Space complexity **O(n)**
- Storing the stego text in a string **STEGO** whose length is n. Space complexity **O(n).**
- Checking the each element of the STEGO and retrieving the secret message of length n. Space complexity **O(n).**
- Changing some elements of SEC to zero according to the value of **array(pos)** . Space complexity **O(n).**

Total space complexity for **Extraction procedure is O(n).**

10 Mathematical formulation for CALP

In this section mathematical analysis of data embedding through pattern changing for various pointed and unpointed letters has been formulated. Figures 27 and 28 show the graphical view of pointed and unpointed letters respectively.

For Pointed Letters

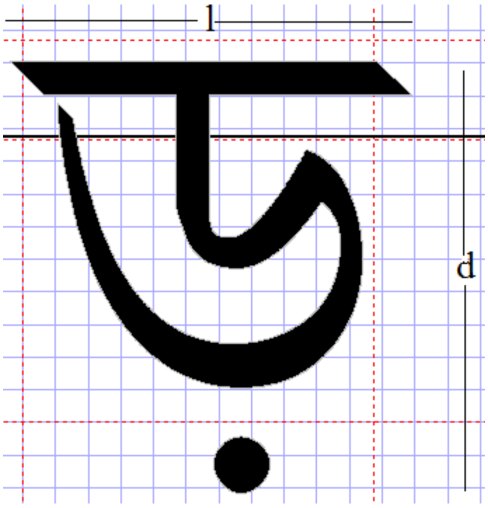


Figure 27: Graphical view of a pointed letter

- Let $x = (d - l) * 5 * (1 + c.s.v)/9$ where c.s.v = current step value and d=height of the letter and l=length of the letter.
- The degree of change of the letter **DC=**$e^{\cosh(x)}$
- For "ড়" $c.s.v = 0$ and For "ঢ়" $c.s.v = 0.5$
- For "র" $c.s.v = 1.0$ and For "য় " $c.s.v = 1.5$

For Unpointed Letters

- Let $x = (d - l) * 5 * (1 + c.s.v)/9$ where c.s.v = current step value and d=height of the letter and l=length of the letter.
- The degree of change of the letter **DC=**$e^{\sinh(x)}$
- For "ত" $c.s.v = 0.3$ and for "ি" $c.s.v = 0.6$
- For" া " $c.s.v = 0.9$ and for "ী" $c.s.v = 1.2$
- For "ষ" $c.s.v = 1.5$ and for "ৎ" $c.s.v = 1.8$
- For"উ" $c.s.v = 2.1$ and for"ক" $c.s.v = 2.4$
- For"ে" $c.s.v = 2.7$ and for"ন" $c.s.v = 3.0$
- For"আ" $c.s.v = 3.3$

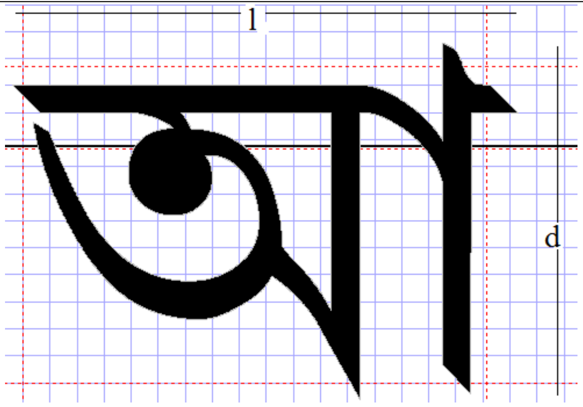


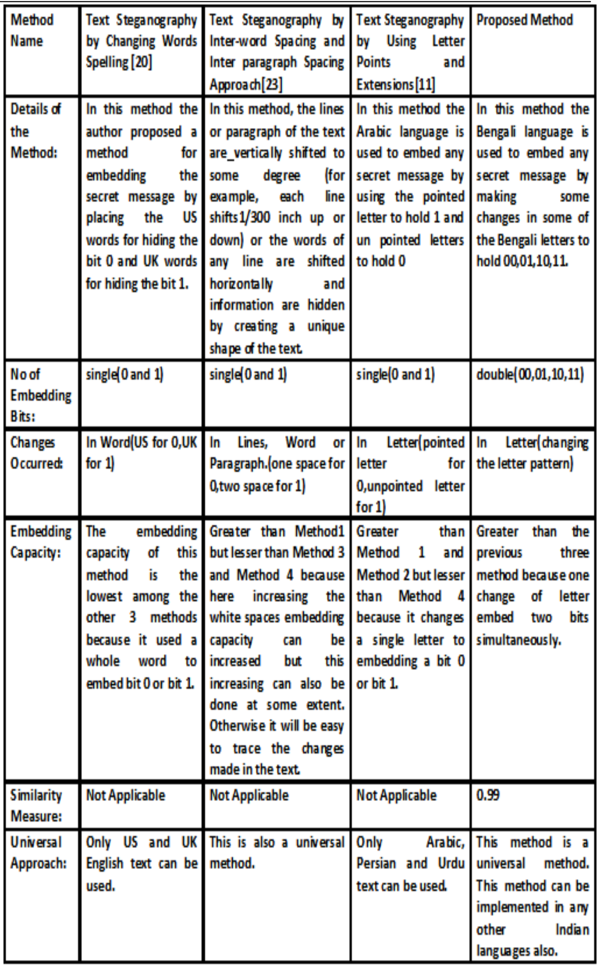
Figure 28: graphical view of an unpointed letter

11 CALP VS Other Text Steganography Methods

In this section a comparison has been shown with some other existing methods like text steganography using Changing word spelling [20],Inter word spacing and inter paragraph spacing [23] or text Steganography by Using Letter Points and Extensions [11]. From the comparative study shown in figure 29 it has been seen that the proposed Bengali Steganography method using CALP is better than the other methods in terms of embedding capacity. This method is an universal one and applicable to any other languages. A technique for measuring the similarity between the cover text and stego text also exist for this method.

12 Concluding Remarks

In this paper the authors presented a novel approach of Bengali text steganography method. Stego text is generated by mapping the binary sequence of the secret message through texture/pattern changes of some alphabets of the cover text in order to achieve high level of security. From figures 25, 26 it has been observed that Single-bit methodology and Double-bit methodology of the proposed method generates the stego text with minimum or zero degradation as both the Jaro score and Correlation-coefficient value is very high. This property also enables the method to avoid the steganalysis. A comparative study with some other existing methods are also shown. The proposed steganography technique through texture/pattern changing is a new approach for the Bengali steganography and this methodology can be extended to any other Indian language also.



Method Name	Text Steganography by Changing Words Spelling [20]	Text Steganography by Inter-word Spacing and Inter paragraph Spacing Approach[23]	Text Steganography by Using Letter Points and Extensions[11]	Proposed Method
Details of the Method:	In this method the author proposed a method for embedding the secret message by placing the US words for hiding the bit 0 and UK words for hiding the bit 1.	In this method, the lines or paragraph of the text are_vertically shifted to some degree (for example, each line shifts1/300 inch up or down) or the words of any line are shifted horizontally and information are hidden by creating a unique shape of the text.	In this method the Arabic language is used to embed any secret message by using the pointed letter to hold 1 and un pointed letters to hold 0	In this method the Bengali language is used to embed any secret message by making some changes in some of the Bengali letters to hold 00,01,10,11.
No of Embedding Bits:	single(0 and 1)	single(0 and 1)	single(0 and 1)	double(00,01,10,11)
Changes Occurred:	In Word(US for 0,UK for 1)	In Lines, Word or Paragraph.(one space for 0,two space for 1)	In Letter(pointed letter for 0,unpointed letter for 1)	In Letter(changing the letter pattern)
Embedding Capacity:	The embedding capacity of this method is the lowest among the other 3 methods because it used a whole word to embed bit 0 or bit 1.	Greater than Method1 but lesser than Method 3 and Method 4 because here increasing the white spaces embedding capacity can be increased but this increasing can also be done at some extent. Otherwise it will be easy to trace the changes made in the text.	Greater than Method 1 and Method 2 but lesser than Method 4 because it changes a single letter to embedding a bit 0 or bit 1.	Greater than the previous three method because one change of letter embed two bits simultaneously.
Similarity Measure:	Not Applicable	Not Applicable	Not Applicable	0.99
Universal Approach:	Only US and UK English text can be used.	This is also a universal method.	Only Arabic, Persian and Urdu text can be used.	This method is a universal method. This method can be implemented in any other Indian languages also.

Figure 29: Comparison between Proposed Method with Other existing Method

References

[1] Alattar, A. and Alattar, O. Watermarking electronic text documents containing justified paragraphs and irregular line spacing. In *Proceedings of SPIE - Security, Steganography, and Watermarking of Multimedia Contents VI*, volume 5306, pages 685–695, June 2004.

[2] Anderson., R. J. and A.P.Petitcolas., F. On the limits of steganography. *IEEE Journal on Selected Areas in Communications (J-SAC), Special Issue on Copyright and Privacy Protection*, 16:474–481, 1998.

[3] Bhattacharyya., S. and Sanyal., G. Study of secure steganography model. In *Proceedings of International Conference on AdvancedComputing and Communication Technologies (ICACCT-2008)*, Panipath,India, 2008.

[4] Bhattacharyya., S. and Sanyal., G. An image based steganography model for promoting global cyber security. In *Proceedings of International Conference on Systemics, Cybernetics and Informatics*, Hyderabad,India, 2009.

[5] Bhattacharyya., S. and Sanyal., G. Implementation and design of an image based steganographic model. In *Proceedings of IEEE International Advance Computing Conference*, Patiala ,India, 2009.

[6] Doerr, G. and Dugelay, J. A guide tour of video watermarking. *Signal Processing: Image Communication*, 18:263–282, 2003.

[7] Doerr, G. and Dugelay, J. Security pitfalls of frameby-frame approaches to video watermarking. *IEEE Transactions on Signal Processing, Supplement on Secure Media*, 52:2955–2964, 2004.

[8] Dowdy, S. and Wearden., S. Statistics for research. *Wiley. ISBN 0471086029.*, 1983.

[9] F. A. P. Petitcolas, R. J. A. and Kuhnl., M. G. Information hiding a survey. In *Proc. of IEEE*, volume 87, pages 1062–1078, 1999.

[10] Gopalan, K. Audio steganography using bit modification. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '03)*, volume 2, pages 421–424, 6-10 April 2003.

[11] Gutub, A. A. and Fattani, M. M. Text steganography by using letter points and extensions. *World Academy of Science, Engineering and Technology 27 ,2007*, pages 13–27, 2007.

[12] Jaro, M. A. Advances in record linking methodology as applied to the 1985 census of tampa florida. *Journal of the American Statistical Society.*, 84:414–420, 1989.

[13] Jaro, M. A. Probabilistic linkage of large public health data file. *Statistics in Medicine 14 (5-7).*, pages 491–498, 1995.

[14] Kahn, D. *The Codebreakers - the comprehensive history of secret communication from ancient times to the Internet*. Scribner, 1996.

[15] Katzenbeisser, S. and Petitcolas., F. A. P. *Information Hiding Techniques for Steganography and Digital Watermarking*. Artech House, London, 2000.

[16] Kevin Curran, K. B. An evaluation of image based steganography methods. *International Journal of Digital Evidence,Fall 2003*, 2003.

[17] L. M. Marvel, J., C. G. Boncelet and Retter., C. T. Spread spectrum image steganography. *IEEE Trans. on Image Processing*, 8:1075–1083, 1999.

[18] lifting scheme, W. S. T. A construction of second generation wavelets. *SIAM J. Math. Anal.*, 29:511–546, 1997.

[19] M. Chapman, G. D. and Rennhard, M. A practical and effective approach to large-scale automated linguistic steganography. In *Proceedings of the Information Security Conference*, pages 156–165, October 2001.

[20] MohammadShirali-Shahreza. Text steganography by changing words spelling. In *ICACT*, 2008.

[21] N. F. Johnson, Z. D. and Jajodia., S. *Information Hiding: Steganography and Digital Watermarking - Attacks and Countermeasures*. Kluwer Academic, 2001.

[22] N.F.Johnson. and Jajodia., S. Steganography: seeing the unseen. *IEEE Computer*, 16:26–34, 1998.

[23] POR, L. Y. and Delina, B. Text steganography by inter-word spacing and inter paragraph spacing approach. *7th WSEAS International Conference on Applied Computer & Applied Computational Science (ACACOS '08), Hangzhou, China,*, April 6-8, 2008.

[24] Prokash, M. A. and Souvik., B. Design of a bengali word peocessor - a new approah. *InfoBiz An Interntional Journal of Informtics,Bussiness & Economics*, 1:13–27, 2008.

[25] R. Calderbank, W. S., I. Daubechies and Yeo., B. Wavelet transforms that map integers to integers. *Applied and Computational Harmonic Analysis.*, 5:332–369, 1998.

[26] R. Chandramouli, N. M. Analysis of lsb based image steganography techniques. In *Proceedings of IEEE ICIP*, 2001.

[27] Shirali-Shahreza, M. and Shahreza, M. S. Text steganography in chat. In *Proceedings of the Third IEEE/IFIP International Conference in Central Asia on Internet the Next Generation of Mobile, Wireless and Optical Communications Networks (ICI 2007)*, Tashkent, Uzbekistan, September 26-28, 2007.

[28] Souvik Bhattacharyya., A. P. K. and Sanyal., G. A novel approach to develop a secure image based steganographic model using integer wavelet transform. In *Proceedings of International Conference on Recent Trends in Information, Telecommunication and Computing (Indexed by IEEE Computer Society)*, Cochin ,India, 2010.

[29] Souvik Bhattacharyya, I. B. and Sanyal, G. A novel approach of secure text based steganography model using word mapping method (wmm). *International Journal of Computer and Information Engineering*, 4:96–103, 2010.

[30] Souvik Bhattacharyya, I. B., Arka Prokash Mazumdar and Sanyal, G. Text steganography using formatting character spacing. *IJICS (International Journal of Information and Computing Science (IJICS))*, 13, 2010.

[31] Souvik Bhattacharyya, S. D., Pabak Indu and Sanyal, G. Hiding data in text through changing in alphabet letter patterns (calp). *Journal of Global Research in Computer Science (JGRCS)*, 2, 2011.

[32] Souvik Bhattacharyya, S. D., Pabak Indu and Sanyal, G. Text steganography using calp with high embedding capacity. *Journal of Global Research in Computer Science (JGRCS)*, 2, 2011.

[33] T Mrkel., J. E. and Olivier., M. An overview of image steganography. In *Proceedings of the fifth annual Information Security South Africa Conference.*, 2005.

[34] W. Bender, N. M., D. Gruhl and Lu., A. Techniques for data hiding. *IBM Systems Journal*, 35:313–316, 1996.

[35] Winkler, W. E. *The state of record linkage and current research problems*. Statistics of Income Division, Internal Revenue Service Publication R99/04., 1999.

[36] Y. Kim, K. M. and Oh, I. A text watermarking algorithm based on word classification and interword space statistics. In *Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR'03)*, pages 775–779, 2003.