SOCIETY FOR SCIENCE AND EDUCATION
UNITED KINGDOM

TMLAI TRANSACTIONS ON MACHINE LEARNING AND ARTIFICIAL INTELLIGENCE

VOLUME 3, ISSUE 3
ISSN 2054-7390

Bilingual Information Hiding System: A Formalized Approach

[1]Riad Jabri, [2]Boran Ibrahim
Faculty of Information Technology, Computer Science Department,
University of Jordan, Amman, Jordan
jabri@ju.eu.jo; boranabed@yahoo.com

ABSTRACT

Steganography and cryptography are used to maintain privacy and security over communication channels. Due to their complexity and diversity, there is a need for their continuous improvements. In this paper, we consider such improvements and propose a new bilingual information hiding system. The proposed system is based on a formal approach that incorporates text-based steganography and cryptography in a way that permits multilayered levels of security and privacy and improves the quality of information hiding in terms of perceptual transparency, robustness and hiding capacity. Considering an Arabic text as a cover one, the inter-word spaces and word features have been used as places for information hiding. As a result, 3.676 have been achieved as an improved average of capacity ratio.

Keywords: Information hiding; Steganography; Cryptography; Encoding; Decoding.

1 Introduction

Steganography is an information hiding technique [1, 2, 3]. For the purpose of our research, we consider steganography as a method of hiding the existence of a bilingual secret message in Arabic text [4]. Text-steganography proceeds according to the following scheme [5]:

- A secret message is concealed in a cover- text using an embedding algorithm to produce a stego- text.
- The stego – text is then transmitted over a communication channel (Internet).
- Upon its delivery, the secret message is recovered using an extracting algorithm.
- The embedding and the extracting algorithms are augmented by the so called a stego- key to encrypt and decrypt the hidden data respectively.

Based on the presented scheme, the secret message is concealed using one of the following methods [5, 6]:

- Modification of the cover-text, such as insertion of spaces, misspelling , modifying the features (name, shape, position, color, size) of the individual characters
- Substitution, such as replacement of insignificant data within the cover text by hidden ones.
- Generation, such as creation of a fake cover.

The most recent efforts, techniques and tools related to our work can be classified as follows:

DOI: 10.14738/tmlai.33.1261
Publication Date: 17th June, 2015
URL: http://dx.doi.org/10.14738/tmlai.33.1261

- Approaches based on inter-word and inter-paragraph spacing to generate dynamic stego-text, as suggested by Por et. al. [7] and Jabri et. al. [5].
- Techniques based on combining the following methods: Open space; syntactic (punctuation) and semantic encoding (synonym words), as suggested by Bender et al. [2].
- Tools based on open space concealing method combined with compression and encryption, such as SNOW as suggested by Kwan [8].
- Technique based on natural language processing and using either the sentence structures or linguistic coding scheme, as suggested by Bergmair [9].
- Techniques specialized for Arabic text, classified into four categories [10]:

1. Dot steganography defined as a vertical displacement of dots in Arabic characters [11].
2. La steganography uses special forms of '' La" word for hiding information [12]
3. Techniques that use letters with extension (kashida) and un-extended ones to hold the secret bits ''one" and ''zero'' respectively [13, 14].
4. Techniques that use a diacritic Arabic text for information hiding [15].

The above-mentioned efforts aim at improving one or more of the following quality indictors: perceptual transparency, robustness and hiding capacity [15]. For example, the hiding algorithms suggested in [13, 14] aim at improving hiding capacity.

Our proposed approach constitutes further improvements of such indictors. In addition, it is characterized by the following:

- The system hides bilingual (Arabic, English) secret messages. It combines inter-word spaces and letter extension (kashida) to hide secret bits. Such combination contributes to perceptual transparency, robustness and increases the hiding capacity.
- The system is based on a formal approach that combines steganography and cryptography as function compositions and introduces multilayered security levels.

The remainder of this paper is organized as follows. Section 2 presents formalization and implementation of the proposed system. Section3 presents analysis and results. A conclusion is given in section 4.

2 The Proposed Method

Based on our set objectives, we define a steganography system in terms of two main stages. The first one is an encoding stage to embed a secret message in a cover text and produce a stego-text. The second one is a decoding stage to extract the secret message from the stego text. The encoding stage transforms the secret message into a sequence of encrypted bits. Such transformation improves the efficiency, the privacy and the security of the hiding process. Then, it considers the cover text as composing of sequence of characters and the uses these characters as positions to hide the secret bits. Such a use is based on the fact that most of Arabic letters can be extended and such extension is considered as one of its writing styles. Thus, the hiding positions are defined as spaces and letters suitable for extension. The spaces are extended by additional space and the letters are extended by kashida (ـ) Hence, the Arabic characters are classified into two sets . The first one includes the Arabic characters that are suitable for extension by kashida. For example, the letters ""ص and ""س are members of such a set. They can be extended by kashida as ـصـ" " and ـسـ" " respectively. The second set includes the special characters (?, : , ", %), the digits (0,..,9) and the non-extendable Arabic characters . For example, the letters"ز" , "و"and" ر" .

Based on such classification, suitability criteria are defined and used to determine the hiding positions in the cover text. The hiding process then proceeds as follows. The secret bits and the cover text are scanned bit by bit and character by character respectively. Upon capturing a hiding position, the respective character (blank or extendable letter) is extended, if the secret bit is “1”. Otherwise, no extension is performed. On the other hand, the decoding stage includes an extracting process to retrieve the hidden bits from their respective positions in the stego-text.

Based on the above-mentioned definition, the encoding and the decoding stages are formalized as a composition of respective functions. The implementation of the proposed system is then reduced to the implementation of an interaction context and the functions from which these stages are composed. They have been implemented according to the algorithms given below using C#.NET as a programming tool. As a result, the proposed system has been constructed with following functionality:

- The sender interacts with the proposed system through an interaction context to facilitate: User authentication; browsing the secret message (SM) and the cover text   (CT) from their respective text files; testing the suitability of CT to hide SM and initiating the encoding process.
- In addition to authentication, the receiver interaction context   facilitates: browsing the stego-text   from its respective text file; and initiating the decoding process.
- The system responds to the sender-initiated request by activating the functions respective to the encoding stage.
- The system responds to the receiver-initiated request by activating the functions respective to the encoding stage.

2.1 The encoding stage

The encoding stage consists of two major steps. First, it transforms the secret message into a sequence of encrypted bits. Second, it considers the cover text as composing of sequence of characters and the uses these characters as positions to hide the secret bits. These steps are formalized and implemented as follows.

Let SM = SM_1 SM_2 … SM_n be the sequence of characters representing the secret message.

The transformation step is defined as a composition of the functions:

$$\text{Bit (Encrypt (Compress (Byte (SM[ ])) } \rightarrow \text{ BSM[ ], where}$$

- Byte (SM[ ]) → SM[ ]   is a function that converts the secret message (SM) to a stream of bytes, using Unicode encoding.
- Compress (SM[ ]) → SM[ ] is a compression function, for example, eliminating extra zero-bytes in the Unicode for Arabic.
- Encrypt (SM[ ]) → ESM[ ] is an encryption function. It uses two passwords of type string to   generate a fixed random stream of keys (Key [ ]). The keys are then XORed with SM [ ] to obtain an encrypted secret message.
- Bit(ESM[ ]) BSM[ ]   is function that converts the encrypted SM into a stream of bits represented as BSM[ ] = BSM_1 BSM_2 … BSM_n.

Let EC = { ت ,…, ب} denotes the set of extendable Arabic characters.

Let NEC = {9…,0,1 ,…. , ر ,ز } denotes the set of non-extendable Arabic characters.

Let PEC = {(ACi, ACj) denotes the set of pair-wise Arabic characters (AC_i, AC_j) such that AC_i EC, AC_j (EC or NEC) and their occurrence as sequence in a text enables the extension of AC_i.

A hiding position is defined as an Arabic character (AC) subject to the following suitability criteria

- Suitability (AC) = True, if (AC (NEC PEC) or AC= " ")
- Suitability (AC) = False, otherwise.

Let CT = CT_1 CT_2 … CT_n is the sequence of characters representing the cover text.

Let HidingPosition (CT[ ]) $\longrightarrow$ HCT[ ] be a function to determine the hiding positions in the cover text.

The embedding process is then defined by the function Embed (BSM, HCT) $\longrightarrow$ ST. However, such embedding is subject to satisfied quality indicators computed by the following functions:

- HidingRatio to express conceivability as (Length (BSM)/Length (hiding positions))%
- CapacityRatio to express conceivability as (Bytes(SM)/Bytes(CT))%

Thus, the encoding stage proceeds according to following steps:

```
Step1: SM = Input (Secret-message);
Step2: BSM[ ] = Bit (Encrypt (Compress (Byte (SM)))
Step3: CT = Select-cover-text
Step4: HCT[ ] = HidingPosition (CT[ ])
Step4:  Compute quality indicators
Step3:  If (Satisfied  (quality indicators){ Embed ( BSM, HCT)}
        Else if { repeat from step3}
```

The implementation of the encoding stage is reduced to the implementation of its functions according to respective algorithms. Representative ones are given below.

Algorithm 1: Encryption and Decryption

Input: compressed stream of bytes respective to SM

Output: encrypted or decrypted stream of bytes respective to SM

Method:

```
encryptedMessage [ ] = encrypt (newMessage, password1) ;
decryptedMessage [ ] = decrypt (encryptedMessage, password2) ;
 encrypt (message [ ] , password1)
  {
    return EncryptDecrypt (message, password1) ;
  }
 decrypt (message [ ] , password2)
  {
    return EncryptDecrypt (message, password1) ;
  }
    EncryptDecrypt (message [ ] , password)
    {
      randomNumbers [ ] = { 08, 06, 02 };
      DerivePasswordBytes = dpb (password, randomNumbers) ;
      key [ ] =dpb.GetBytes(128) ;  // Return 128 random numbers
      returnMessage [ message. length ] ;
      for i = 1 To message. length        step 1
       {
         index = i mod key. length ;
         returnMessage [ i ] = key [index]  XOR  message [ i ] ;
       }
      return returnMessage ;
    }
```

Algorithm 1 performs encryption and decryption. It uses two passwords of type string to generate a fixed random stream of keys The keys are then XORed with SM (or CT)to obtain an encrypted (or decrypted) secret message. By using passwords and randomly generated keys, Algorithm 1 introduces two levels of security.

Algorithm 2: Hiding positions

Input: Covertext (CT)

Output: Hiding positions in CT.

Method:

```
Hiding positions [ ] = Suitable (CT[ ] )
Suitable (CT[ ] )
 {  for i =1  To  CT. length     step 1
   {
     if ( Suitability (CT [i] = True )  Then
       { HidingPositions  [i ] = " 1"
       Else
        {  HidingPositions  [i ] = " 0"
     }
```

Algorithm 3: Embedding

Input: The stream of bits BSM respective to the secret message SM.

The over text (CT)

The hiding positions (HidingPositions) in CT

Output: stego_text (secret bits inside cover text).

Method:

```
Embed  (BSM , CT, HidingPositions )
{
  cover = CT ;
  result = "  " ;
  coverIndex = 1 ;
  for i =1  To  message. length     step 1
    {
      while ( not HidingPositions [coverIndex ] = "1 ")
       Do
        { result = result + cover [ coverIndex ] ;
          coverIndex = coverIndex + 1
        }
        if CT [coverIndex ] = "  " )   Then
           {
             result = result + cover [ coverIndex ] + "  " ;
             coverIndex = coverIndex + 1 ;
           }
          else
           {
            result = result + cover [ coverIndex ] + " - " ;
            coverIndex = coverIndex + 1 ;
     }
     Insert an end marker for BSM at cover [ coverIndex ]
     Complete the stego text by remaining cover text:
      for i = coverIndex + 1  To  cover. length   step 1
        result = result + cover [ i ] ;
    return result ;
}.
```

2.2 The Decoding Stage

The decoding stage extracts the secret message from the stego-text at receiver's side according to the following steps:

Step 1: ST = Browse (stego-text)

Step 2: If (authentication = True) {Decompress (Decrypt (Byte (Bit-extract (ST)))}, where

- Bit-extract (ST) BSM[ ] is function that extracts the encrypted SM as the stream of bits BSM[ ] = BSM_1 BSM_2 … BSM_n
- Byte (BSM[ ]) SM[ ] is a function that converts extracted message (BSM) to a stream of bytes, represented as SM[ ]
- Decrypt (SM[ ]) SM[ ] XOR Key [ ] is an decryption function. As in Encrypt, the function Decrypt uses two passwords of type string to generate a fixed random stream of keys. However, the keys are XORed with SM [ ] to obtain a decrypted secret message.
- Decompress (SM[ ]) SM[ ] is a decompression function to recover SM as represented by Unicode for Arabic.

The implementation of the decoding stage is reduced to an interaction context for authentication, browsing the stego text and then displaying the secret message. This is achieved by applying the functions respective to the decoding stage. These functions have been implemented according to respective algorithms. Representative ones are given below.

Algorithm 4: Bit-Extract

Input: stego_text

Output: stream of bits

Method:

```
lastPosition = message.LastIndexOf ("   ") ;    // 3 spaces
if ( lastPosition < 0 )   Then
  {
     lastPosition = getLastPosition ( message ) ;
  }
temp = message ;
bits = extractBits ( lastPosition , temp ) ;
getLastPosition ( message )
  {
    for i = 1   to   message. length     step 1
        if ( message[ i ] = ' - ' AND  message [ i +1] = ' - ' )   Then
            return i +1 ;
    return i ;
  }
extractBits ( lastPosition , temp )
{
bits = "  " ;
for i = 0   To   lastPosition    step 1
   if ( temp [ i ] = ' ' )    Then
    {
        if ( temp [ i +1] = ' ' )    Then
          {
            bits = bits + " 1 " ;
            i = i + 1 ;
          }
        else
            bits = bits + " 0 " ;
    }
   else if ( temp [ i ] = ' - ' )   Then
    {
        bits = bits + " 1 " ;
    }
   else
    {
        if ( temp [ i + 1 ] <> ' - ' )    Then
           if checkTwoCharacters ( temp [ i ] , temp [ i + 1 ] ) )   Then
            {
                bits = bits + " 0 " ;
            }
    }
    return bits ;
}.
```

Algorithm 5: Conversion to bytes

Input: stream of bits

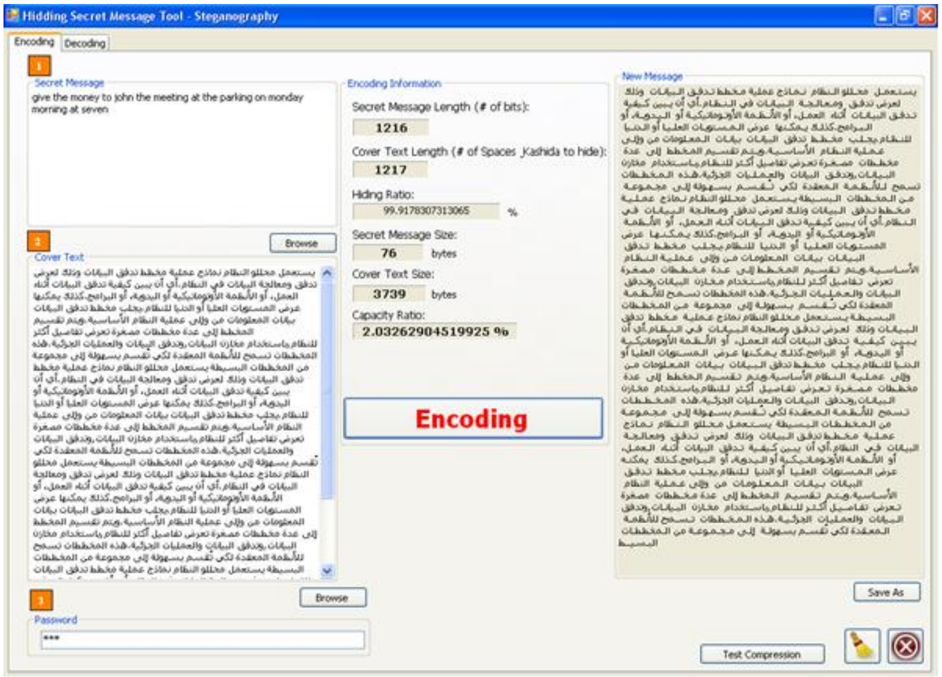
Output: stream of bytes

Method:

```
message [ ] = decode (bits) ;
  decode (bits)
  {
     hidden [bits. length / 8] ;
     c = 0 ;
     for i =1 To hidden. length      step 1
       for j = 7 To 0    step -1
         {
            if (bits [c++] = '1' )  Then
            hidden [ i ] = (hidden [ i ] OR (1<< j)) ;
         }
     return hidden ;
  }.
```

3 Experiments and Analysis

The proposed system has been developed with an interface represented by two forms. The first form is denoted facilitates interaction with the presented encoding stage and its respective concealing functions. Furthermore, it displays quality indicators such as hiding and capacity ratios respective to the secret message and the browsed cover text. The second form facilitates interaction with the presented decoding stage and its respective functions. Through its interface, the proposed steganography system has been tested using several multilingual texts (Arabic and English) as demonstrated by Figure 1, Figure 2, Figure 3 and Figure 4 respectively.



Secret Message: give the money to john the meeting at the parking on monday morning at seven

Cover Text: يستعمل محللو النظام نماذج عملية مخطط تدفق البيانات وذلك لعرض تدفق ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو البرامج.كذلك يمكنها عرض المستويات العليا أو الدنيا للنظام بجلب مخطط تدفق البيانات بيانات المعلومات من وإلى عملية النظام الأساسية.ويتم تقسيم المخطط إلى عدة مخططات مصغرة تعرض تفاصيل أكثر للنظام باستخدام مخازن البيانات,وتدفق البيانات والعمليات الجزئية.هذه المخططات تسمح للأنظمة المعقدة لكي تقسم بسهولة إلى مجموعة من المخططات البسيطة يستعمل محللو النظام نماذج عملية مخطط تدفق البيانات وذلك لعرض تدفق ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو البرامج.كذلك يمكنها عرض المستويات العليا أو الدنيا للنظام بجلب مخطط تدفق البيانات بيانات المعلومات من وإلى عملية النظام الأساسية.ويتم تقسيم المخطط إلى عدة مخططات مصغرة تعرض تفاصيل أكثر للنظام باستخدام مخازن البيانات,وتدفق البيانات والعمليات الجزئية.هذه المخططات تسمح للأنظمة المعقدة لكي تقسم بسهولة إلى مجموعة من المخططات البسيطة يستعمل محللو النظام نماذج عملية مخطط تدفق البيانات

Password: ***

Encoding Information

Secret Message Length (# of bits): 1216

Cover Text Length (# of Spaces _Kashida to hide): 1217

Hiding Ratio: 99.9178307313065 %

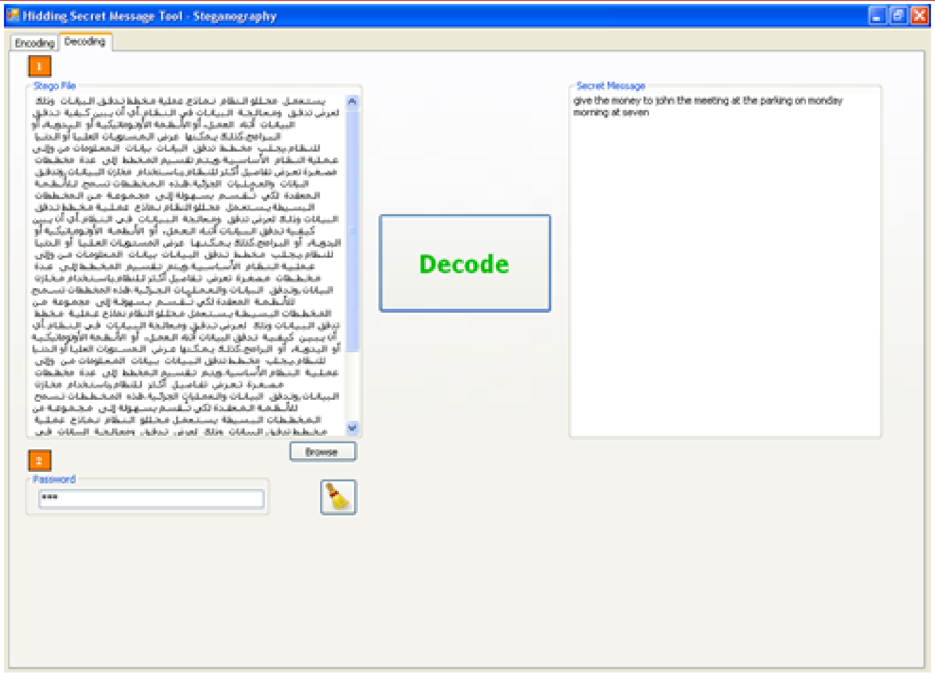
Secret Message Size: 76 bytes

Cover Text Size: 3739 bytes

Capacity Ratio: 2.03262904519925 %

New Message: يستعمل محللو النظام نماذج عملية مخطط تدفق البيانات وذلك لعرض تدفق ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو البرامج.كذلك يمكنها عرض المستويات العليا أو الدنيا للنظام بجلب مخطط تدفق البيانات بيانات المعلومات من وإلى عملية النظام الأساسية.ويتم تقسيم المخطط إلى عدة مخططات مصغرة تعرض تفاصيل أكثر للنظام باستخدام مخازن البيانات,وتدفق البيانات والعمليات الجزئية.هذه المخططات تسمح للأنظمة المعقدة لكي تقسم بسهولة إلى مجموعة من المخططات البسيطة يستعمل محللو النظام نماذج عملية مخطط تدفق البيانات وذلك لعرض تدفق ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو البرامج.كذلك يمكنها عرض المستويات العليا أو الدنيا للنظام بجلب مخطط تدفق البيانات بيانات المعلومات من وإلى عملية النظام الأساسية.ويتم تقسيم المخطط إلى عدة مخططات مصغرة تعرض تفاصيل أكثر للنظام باستخدام مخازن البيانات,وتدفق البيانات والعمليات الجزئية.هذه المخططات تسمح للأنظمة المعقدة لكي تقسم بسهولة إلى مجموعة من المخططات البسيطة يستعمل محللو النظام نماذج عملية مخطط تدفق البيانات وذلك لعرض تدفق ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو البرامج.كذلك يمكنها عرض المستويات العليا أو الدنيا للنظام بجلب مخطط تدفق البيانات بيانات المعلومات من وإلى عملية النظام الأساسية.ويتم تقسيم المخطط إلى عدة مخططات مصغرة تعرض تفاصيل أكثر للنظام باستخدام مخازن البيانات,وتدفق البيانات والعمليات الجزئية.هذه المخططات تسمح للأنظمة المعقدة لكي تقسم بسهولة إلى مجموعة من المخططات البسيط

Figure 1: Encoding of English text

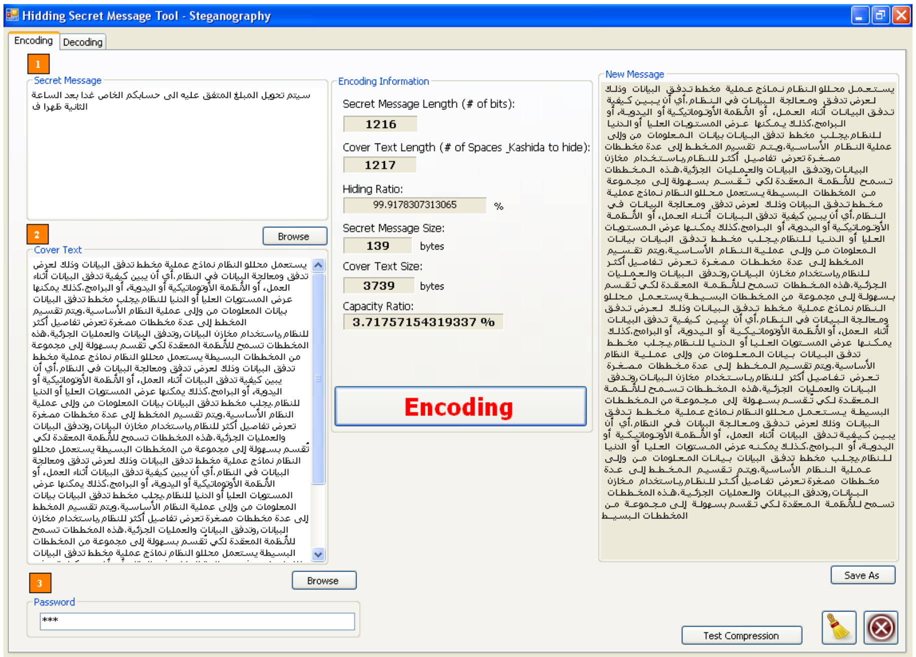


يستعمل محللو النظام نماذج عملية مخطط تدفق البيانات وذلك
لعرض تدفق ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق
البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو
البرامج.كذلك يمكنها عرض المستويات العليا أو الدنيا
للنظام.يجلب مخطط تدفق البيانات بيانات المعلومات من وإلى
عملية النظام الأساسية.ويتم تقسيم المخطط إلى عدة مخططات
مصغرة تعرض تفاصيل أكثر للنظام,باستخدام مخازن البيانات,وتدفق
البيانات والعمليات الجزئية.هذه المخططات تسمح للأنظمة
المعقدة لكي تقسم بسهولة إلى مجموعة من المخططات
البسيطة يستعمل محللو النظام نماذج عملية مخطط تدفق
البيانات وذلك لعرض تدفق ومعالجة البيانات في النظام.أي أن يبين
كيفية تدفق البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية أو
اليدوية، أو البرامج.كذلك يمكنها عرض المستويات العليا أو الدنيا
للنظام.يجلب مخطط تدفق البيانات بيانات المعلومات من وإلى
عملية النظام الأساسية.ويتم تقسيم المخطط إلى عدة
مخططات مصغرة تعرض تفاصيل أكثر للنظام,باستخدام مخازن
البيانات,وتدفق البيانات والعمليات الجزئية.هذه المخططات تسمح
للأنظمة المعقدة لكي تقسم بسهولة إلى مجموعة من
المخططات البسيطة يستعمل محللو النظام نماذج عملية مخطط
تدفق البيانات وذلك لعرض تدفق ومعالجة البيانات في النظام.أي
أن يبين كيفية تدفق البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية
أو اليدوية، أو البرامج.كذلك يمكنها عرض المستويات العليا أو الدنيا
للنظام.يجلب مخطط تدفق البيانات بيانات المعلومات من وإلى
عملية النظام الأساسية.ويتم تقسيم المخطط إلى عدة مخططات
مصغرة تعرض تفاصيل أكثر للنظام,باستخدام مخازن
البيانات,وتدفق البيانات والعمليات الجزئية.هذه المخططات تسمح
للأنظمة المعقدة لكي تقسم بسهولة إلى مجموعة من
المخططات البسيطة يستعمل محللو النظام نماذج عملية
مخطط تدفق البيانات وذلك لعرض تدفق ومعالجة البيانات في

Password: ***

Secret Message: give the money to john the meeting at the parking on monday morning at seven

Figure 2: Decoding of English text



Secret Message

سيتم تحويل المبلغ المتفق عليه الى حسابكم الخاص غدا بعد الساعة
الثانية ظهرا ف

Cover Text

يستعمل محللو النظام نماذج عملية مخطط تدفق البيانات وذلك لعرض
تدفق ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق البيانات أثناء
العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو البرامج.كذلك يمكنها
عرض المستويات العليا أو الدنيا للنظام.يجلب مخطط تدفق البيانات
بيانات المعلومات من وإلى عملية النظام الأساسية.ويتم تقسيم
المخطط إلى عدة مخططات مصغرة تعرض تفاصيل أكثر
للنظام,باستخدام مخازن البيانات,وتدفق البيانات والعمليات الجزئية.هذه
المخططات تسمح للأنظمة المعقدة لكي تقسم بسهولة إلى مجموعة
من المخططات البسيطة يستعمل محللو النظام نماذج عملية مخطط
تدفق البيانات وذلك لعرض تدفق ومعالجة البيانات في النظام.أي أن
يبين كيفية تدفق البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية أو
اليدوية، أو البرامج.كذلك يمكنها عرض المستويات العليا أو الدنيا
للنظام.يجلب مخطط تدفق البيانات بيانات المعلومات من وإلى عملية
النظام الأساسية.ويتم تقسيم المخطط إلى عدة مخططات مصغرة
تعرض تفاصيل أكثر للنظام,باستخدام مخازن البيانات,وتدفق البيانات
والعمليات الجزئية.هذه المخططات تسمح للأنظمة المعقدة لكي
تقسم بسهولة إلى مجموعة من المخططات البسيطة يستعمل محللو
النظام نماذج عملية مخطط تدفق البيانات وذلك لعرض تدفق ومعالجة
البيانات في النظام.أي أن يبين كيفية تدفق البيانات أثناء العمل، أو
الأنظمة الأوتوماتيكية أو اليدوية، أو البرامج.كذلك يمكنها عرض
المستويات العليا أو الدنيا للنظام.يجلب مخطط تدفق البيانات بيانات
المعلومات من وإلى عملية النظام الأساسية.ويتم تقسيم المخطط
إلى عدة مخططات مصغرة تعرض تفاصيل أكثر للنظام,باستخدام مخازن
البيانات,وتدفق البيانات والعمليات الجزئية.هذه المخططات تسمح
للأنظمة المعقدة لكي تقسم بسهولة إلى مجموعة من المخططات
البسيطة يستعمل محللو النظام نماذج عملية مخطط تدفق البيانات

Password: ***

Encoding Information

Secret Message Length (# of bits): 1216

Cover Text Length (# of Spaces _Kashida to hide): 1217

Hiding Ratio: 99.9178307313065 %

Secret Message Size: 139 bytes

Cover Text Size: 3739 bytes

Capacity Ratio: 3.71757154319337 %

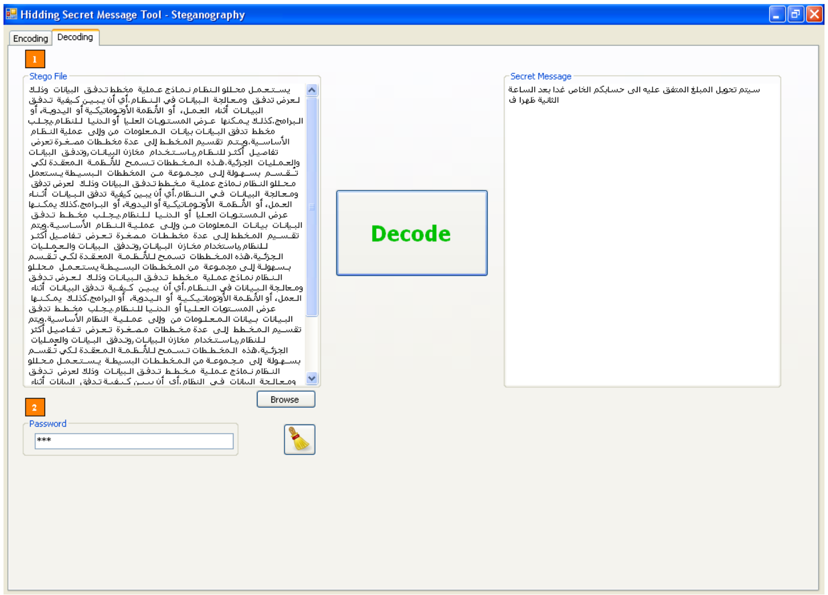
New Message

يستعمل محللو النظام نماذج عملية مخطط تدفق البيانات وذلك
لعرض تدفق ومعالجة البيانات في النظام.أي أن يبين كيفية
تدفق البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو
البرامج.كذلك يمكنها عرض المستويات العليا أو الدنيا
للنظام.يجلب مخطط تدفق البيانات بيانات المعلومات من وإلى
عملية النظام الأساسية.ويتم تقسيم المخطط إلى عدة مخططات
مصغرة تعرض تفاصيل أكثر للنظام,باستخدام مخازن
البيانات,وتدفق البيانات والعمليات الجزئية.هذه المخططات
تسمح للأنظمة المعقدة لكي تقسم بسهولة إلى مجموعة
من المخططات البسيطة يستعمل محللو النظام نماذج عملية
مخطط تدفق البيانات وذلك لعرض تدفق ومعالجة البيانات في
النظام.أي أن يبين كيفية تدفق البيانات أثناء العمل، أو الأنظمة
الأوتوماتيكية أو اليدوية، أو البرامج.كذلك يمكنها عرض المستويات
العليا أو الدنيا للنظام.يجلب مخطط تدفق البيانات بيانات
المعلومات من وإلى عملية النظام الأساسية.ويتم تقسيم
المخطط إلى عدة مخططات مصغرة تعرض تفاصيل أكثر
للنظام,باستخدام مخازن البيانات,وتدفق البيانات والعمليات
الجزئية.هذه المخططات تسمح للأنظمة المعقدة لكي تقسم
بسهولة إلى مجموعة من المخططات البسيطة يستعمل محللو
النظام نماذج عملية مخطط تدفق البيانات وذلك لعرض تدفق
ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق البيانات
أثناء العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو البرامج.كذلك
يمكنها عرض المستويات العليا أو الدنيا للنظام.يجلب مخطط
تدفق البيانات بيانات المعلومات من وإلى عملية النظام
الأساسية.ويتم تقسيم المخطط إلى عدة مخططات مصغرة
تعرض تفاصيل أكثر للنظام,باستخدام مخازن البيانات,وتدفق
البيانات والعمليات الجزئية.هذه المخططات تسمح للأنظمة
المعقدة لكي تقسم بسهولة إلى مجموعة من المخططات
البسيطة يستعمل محللو النظام نماذج عملية مخطط تدفق
البيانات وذلك لعرض تدفق ومعالجة البيانات في النظام.أي أن
يبين كيفية تدفق البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية أو
اليدوية، أو البرامج.كذلك يمكنه عرض المستويات العليا أو الدنيا
للنظام.يجلب مخطط تدفق البيانات بيانات المعلومات من وإلى
عملية النظام الأساسية.ويتم تقسيم المخطط إلى عدة
مخططات مصغرة تعرض تفاصيل أكثر للنظام,باستخدام مخازن
البيانات,وتدفق البيانات والعمليات الجزئية.هذه المخططات
تسمح للأنظمة المعقدة لكي تقسم بسهولة إلى مجموعة من
المخططات البسيط

Figure 3: Decoding of an Arabic text

Table 1 shows the achieved capacity ratios for secrete messages and cover texts of different size, where 3.67 is indicated as the average capacity ratio. This ratio is higher than the ones reported by other approaches as demonstrated by table 2. In [1], 2.46 and3.73 are reported as the average capacity ratio using one and three kashidas respectively. However, our approach is based on using one kashida per letter. Using the same text as a cover one, Table 3 shows the size of the stego_text for Arabic and

English secrete messages. In both cases the size of cover text has been increased by 27.6% and 29% respectively.



Stego File:

يستعمل محللو النظام نماذج عملية مخطط تدفق البيانات وذلك
لعرض تدفق ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق
البيانات أثناء العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو
البرامج.كذلك يمكنها عرض المستويات العليا أو الدنيا للنظام.يجلب
مخطط تدفق البيانات بيانات المعلومات من وإلى عملية النظام
الأساسية.ويتم تقسيم المخطط إلى عدة مخططات مصغرة تعرض
تفاصيل أكثر للنظام,باستخدام مخازن البيانات,وتدفق البيانات
والعمليات الجزئية.هذه المخططات تسمح للأنظمة المعقدة لكي
تقسم بسهولة إلى مجموعة من المخططات البسيطة يستعمل
محللو النظام نماذج عملية مخطط تدفق البيانات وذلك لعرض تدفق
ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق البيانات أثناء
العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو البرامج.كذلك يمكنها
عرض المستويات العليا أو الدنيا للنظام.يجلب مخطط تدفق
البيانات بيانات المعلومات من وإلى عملية النظام الأساسية.ويتم
تقسيم المخطط إلى عدة مخططات مصغرة تعرض تفاصيل أكثر
للنظام,باستخدام مخازن البيانات,وتدفق البيانات والعمليات
الجزئية.هذه المخططات تسمح للأنظمة المعقدة لكي تقسم
بسهولة إلى مجموعة من المخططات البسيطة يستعمل محللو
النظام نماذج عملية مخطط تدفق البيانات وذلك لعرض تدفق
ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق البيانات أثناء
العمل، أو الأنظمة الأوتوماتيكية أو اليدوية، أو البرامج.كذلك يمكنها
عرض المستويات العليا أو الدنيا للنظام.يجلب مخطط تدفق
البيانات بيانات المعلومات من وإلى عملية النظام الأساسية.ويتم
تقسيم المخطط إلى عدة مخططات مصغرة تعرض تفاصيل أكثر
للنظام,باستخدام مخازن البيانات,وتدفق البيانات والعمليات
الجزئية.هذه المخططات تسمح للأنظمة المعقدة لكي تقسم
بسهولة إلى مجموعة من المخططات البسيطة يستعمل محللو
النظام نماذج عملية مخطط تدفق البيانات وذلك لعرض تدفق
ومعالجة البيانات في النظام.أي أن يبين كيفية تدفق البيانات أثناء

Password: ***

Secret Message:

سيتم تحويل المبلغ المتفق عليه الى حسابكم الخاص غدا بعد الساعة
الثانية ظهرا ف

Figure 4: Decoding of an Arabic text

Table 1: The achieved capacity ratio

Cover size (byte)	Number of hidden bits	Capacity ratio (%)	Average capacity (%)
1714	560	3.676	3.657
6976	2288	3.655	
19264	6320	3.649	
31552	10352	3.648	

Table 2: The capacity ratios achieved by other approaches

Approach	Capacity ratio (%)
Dots[15]	1.37
Kashida [13]	2.46-3.73
Kashida [14]	3.09
Diactrics [15]	3.27

Table 3: The size of the secret message

Type of the secret message	Size of the cover text	Size of the stego_text
Arabic	3.65 kb	4.66 kb
English	3.65 kb	4.71 kb

In addition to its efficiency and flexibility, the proposed text steganography system has demonstrated improvements in its quality indicators as follows:

- The adopted encoding methods are characterized by their flexibility and enables hiding dynamic secret messages. The binary conversion and the encryption constitute two level of security
- The adopted embedding criteria combine suitability and randomness to ensure robustness of the stego-text, perceptual transparency and improved hiding capacity.

Compared to similar systems, our system has a higher increase in the size of the cover text. For example, the average size according to our system is 25% and the one according to the system proposed in [5] is 2%. However, the proposed system has better results in terms of the following:

- A higher utilization percentage of the cover- text. This demonstrated by the ratio between the size of the secret message and the one of the cover text. For example, our approach has achieved 1:3 ratio while in [7] a cover text with a size < 16KB is required for a secret message with a length < 4 KB.
- A higher hiding capacity as demonstrated by Table 1 and Table 2.

4 Conclusion

In this research, an information hiding model has been suggested. Based on such model, a text steganography system has been implemented. The system is characterized by its scalability and flexibility. Although the proposed system has better quality indicators than the ones for similar system, more improvements are needed for such indicators. Mainly, the capacity of the cover- text and the robustness of the stego- text. Hence, efforts in this direction constitute a future work.

REFERENCES

[1]. Bennett, K., *Linguistic steganography: Survey, analysis and robustness concerns for information hiding in text*. Purdue University, CERIAS Technical Report, 2004.

[2]. Bender, S., Gruhl, D., Morimoto, N., *Techniques for data hiding.* IBM System Journal, 1996. 35(4): p. 313-336.

[3]. Komal, P., Sumit, V., Hitesh, C., *A Survey of Information Hiding Techniques*. International Journal of Emerging Technology and Advanced Engineering, 2013. 3(1): p.347-349

[4]. Cachin, C., *An information –theoretic model for steganography*. Lecture Notes in Computer Science, 1998. 1225: p. 306-318.

[5]. Jabri, R., Ibrahim, B., Al-Zoubi, H., *Information Hiding: A Generic Approach*. Journal of Computer Science, 2009. 5(12): 930-936.

[6]. Sumathi, G., Santanam, T., Umamaheswari, G., *A study of Various steganogr762arphic Techniques Used for Information Hiding.* International Journal of Emerging Technology and Advanced Engineering, 2013. 4(6): P. 9-25.

[7]. Por, L., Delina, B., *Information hiding: A new approach in text steganography*. 7th WSEAS Int. Conf. on Applied Computer & Applied Computational Science, Hanghou, Chiana, 2008. P. 689-695.

[8]. Kwan, M., The SNOW homepage, http://www.darkside.com.au/snow/, 1998.

[9]. Haddouch Bergmair, R., *Towards linguistics stganogrphy: A systematic investigation of approaches, Systems and Issues*. Technical Report. 2004.

[10]. Shirali-Shareza, M.,H., Shirali-Shareza, M. *Steganography in Persian and Arabic Unicode Texts Using Pseudo-Space and Pseudo-Connection Characters*. Journal of Theoretical and Applied Information, 2008. 4(8): p. 682-687.

[11]. Shirali-Shareza, M., H., Shirali-Shareza, M. *A New Approach to Persian/Arabic Text Steganography*. 5th IEEE/ACIS on Computer and Information Science (ICIS- COMSAR 06), 2006. P. 310-315.

[12]. Shirali-Shareza, M. *A New Persian/Arabic Text Stegonography Using " La Word"*. Proceedings of the International Joint Conference on Computer**,** Information and Systems Sciences and Engineering**,** Bridgeport, CT, USA, 2007.

[13]. Al Haidari, F., Gutub, A., Al-Kahsah, K., Hamodi, J. 2009. *Improved Security and Capacity for Arabic Text Steganography Using 'Kashida ' Extensions*. IEEE/ACS International Conference on Computer Systems and Applications, Rabat, 2009. P. 396-399.

[14]. Al-Azawi, A.F., Fadhil, M.,A., *Arabic Text Steganography Using Kashida Extensions with Huffman Code*. Journal of Applied Sciences 2010. 10(5): P. 436-439.

[15]. Gutub, A., Elarian, Y., Awaideh, S., Alvi, A., *Arabic Text Steganography Using Multiple Diacritics*, WoSPA 5th IEEE International Workshop on Signal Processing and its Applications, Sharjah, UAE, 2008.