Journal of Advanced Sciences and Engineering Technologies (2023) 5 (2): 49-57
https://doi.org/10.32441/jaset.05.02.05

Journal of Advanced Sciences and Engineering Technologies
https://isnra.net/ojs/index.php/jaset/index/

The use of binary digit mapping on ASCII characters to create a high-capacity, undetectable text steganography.

Eng. Saad Nasser Al Azzam

Dr. Fahad Garni

College of Computing - Department of Cyber Security - University of Bisha

Kingdom Saudi Arabia

Snazzam.199@gmail.com

Keywords:

Artificial Intelligent, Smart home, GSM Network

ARTICLE INFO

Article history:

Received	03 February. 2023
Accepted	10 March. 2023
Available online	02 April.2023

Citation : Al.Azzam , S. N., & Al-Garni, F. A. . (2023). The use of binary digit mapping on ASCII characters to create a high-capacity, undetectable text steganography. Journal of Advanced Sciences and Engineering Technologies, 5(2), 49–57.
https://doi.org/10.32441/jaset.05.02.05

ABSTRACT

Plain text is commonly utilized for online news and social media transmission due to its lightweight and cross-platform nature. Unfortunately, it may easily be exploited by attackers (For instance, in cases when the language is altered for specific goals, illegal access to the material or its abuse is possible.) To solve this problem, we embed secret texts (ST) within seemingly innocuous cover texts (CT) so that we can track CT for modifications. Based on the findings of this study, a method is provided for ST embedding in CT in which binary ST digits are converted to CT binary digits through Characters in ASCII (containing whitespace, punctuation, and special characters). Before commencing the embedding procedure, Using a One-Time Pad (OTP), the ST text was translated into cipher text, and the value of each character was changed into a binary integer with a length of seven bits. As opposed to the ST text, which required additional time to encode, the CT text could be converted into a binary integer using just the first seven bits. During the embedding process, It was decided to swap the first bit of the ST character for the first bit of the CT character, which both had the same amount of bits (putting the first bit of the ST character, for example, on the same line as the first bit of the CT character). Any piece of ST that was recorded on any bit of CT served as a stage key. Extracting ST from CT requires the stage key. The embedded information served effectively as text steganography or watermark, and the CT's appearance was unaffected by the embedding procedure. In other words, the 1 character of ST might be concealed by using 7 CT characters. Also, a Jaro-Winkler Distance of 1 meant that the stage texts made couldn't be told apart from CT by looking at them.

[*] Corresponding Author Eng. Saad Nasser Al Azzam, College of Computing - Department of Cyber Security - University of Bisha, Kingdom Saudi Arabia, Snazzam.199@gmail.com

Introduction

Online news, social media, and other online sources are becoming increasingly available to a growing number of people all over the world. This is especially true for those who have access to a computer.

Smartphones or other devices. Most of the data found in these virtual communities is stored in an easily-modified plain-text format. Because the original was changed without good reason, the result is false information or a hoax. Steganography, also known as the watermark technique, is used to safeguard information by concealing secret content behind a seemingly innocuous cover text. This is essential since it is very challenging to maintain the information's validity and integrity when it is presented in plain text. This study suggests a technique for embedding a hidden message in ordinary ASCII text by mapping the binary value of the letters. Using this strategy, the plain text that has the secret text inserted into it will always seem very similar to the plain text that does not include the secret information. This is a guarantee. If an unauthorized user attempts to change the plain text, either by deleting or adding characters, this will be instantly visible. As a result, the plain text containing the secret material may be believed to be genuine and original. This method may be used to any plain-text digital information, such as online news, e-commerce, social networking, and so on. It's useful for SMS, too (SMS). Information concealment, watermarking, and steganography are the three most common uses for data embedding. Despite their apparent conceptual and technological closeness, these three topics have significant differences [1, 2, 3, and 4].

Watermarks are used to hide information in media where the emphasis is on the watermark, whereas steganography is used to hide secret information in media when the attention is on the material itself as the primary focus of the medium.

Thus, steganography is successful only if the intended recipient never discovers the concealed information [5, 6, and 7].

The term "steganography," which refers to the practice of concealing written information, dates back to antiquity. A slave called according to Herodotus [8, 9]. Was the first person to utilize steganography His master had hidden instructions in a tattoo on his skull. The slave was sent to the Miletus, the capital of the Ionians, as soon as his hair started to grow again so that he could report to King Aristagoras.

This secret message was then revealed by shaving the head of the slave. The modern practice of steganography involves concealing a message inside another file, which may take the shape of text, photos, audio, video, or any other type of digital information [10,11,12].

The term "stage file" refers to the file that contains the data that has been embedded, "stage key" refers to the key that was used to encode the process of hiding the data, and "embedded data" refers to the data that has been hidden. Text steganography refers to the practice of utilizing text to obfuscate sensitive information. It uses a number of techniques (such as changing the text's structure or vocabulary or rearranging the random letters or using context-free grammar) to generate comprehensible writing [13,14,15].

Robust and fragile steganography are defined by their respective strengths and weaknesses. The goal of robust steganography is to stop anyone from finding the hidden data inside the file, while the goal of fragile steganography is to stop anyone from changing the stego file in any way [13.14.15.].

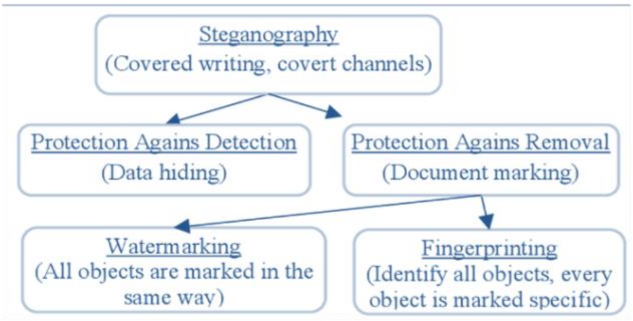

Steganography
(Covered writing, covert channels)
Protection Agains Detection
(Data hiding)
Protection Agains Removal
(Document marking)
Watermarking
(All objects are marked in the same way)
Fingerprinting
(Identify all objects, every object is marked specific)

Figure 1 Directions within steganography [1].

This study is made up of several parts: an introduction, a discussion of relevant literature and theories, a description of the method to be used, the results and discussion, and finally a summary and conclusion.

2. Related Works and Theories

Using text steganography as an example, we'll go through the reasoning behind our suggested technique of data embedding.

A. Hiding Data in Wordlist

Using dynamic CT, which is explained in detail in the method that is referred to as "Hiding Data in Wordlist," it is possible to hide ST inside a word list. Calculating the decimal value of each secret character results in the production of the first character, which is then encoded in the first letter and a word of a preset length. This first character is acquired by calculating the decimal value of each secret character. [16].

B. Using Paragraphs to Hiding Secret Text

To conceal ST, paragraphs of English text are used as CT in text steganography [17] It generates a stego key that may be used to encode and decode hidden data. In appearance, the stego key generated is identical to CT. The approach creates a bit stream by using the binary representations of ASCII characters. This bit stream is then used in conjunction with a One Time Pad in order to encrypt the secret text that is inserted (OTP). The chipped text is painstakingly reassembled into the CT words in the correct order.Depending on the term, either the first or final letter of the word is modified by inserting a bit of 0 or 1. But it doesn't work for words whose first and last letters are the same. Based on prior studies' findings, In order to conceal information, The approach of "Hiding Data in Paragraphs" was developed [8]. Using the binary value of letters and the value of the XOR between the first and final words of a line in a CT paragraph, this technique hides data. The Hiding a Bit of ST in Each Character in the Phrases of a CT approach was developed as a result of other studies [10], and it involves disguising little amounts of ST inside the letters that make up the words of a CT. Selecting a certain CT character is used to mask each individual bit.

C. One Time Pad Encryption

To encrypt and decrypt data, One-Time Pad (OTP) employs a single key $\left\lfloor \frac{\max|s_1|,|s_2|}{2} \right\rfloor - 1$ for both operations. OTP encryption is a popular cryptographic method. OTP is very difficult to break [11] because each key can only be used once and there are as many keys as there are encoded messages. Since the key sequences don't follow any kind of statistical pattern, the attacker will have to try everything before they succeed. The notation used by the OTP algorithm, which employs XOR or modulo (mod), looks like this [18].

$$C_i = (P_i + K_i) \bmod z \qquad (1)$$

If Ci is the ciphertext, Pi denotes the plaintext, Ki denotes the key, and z denotes the total amount of characters that may be used,

D. Plain Text

Unlike other forms of text, plain text does not specify details like font family, size, or colour or layout. In the absence of agreement, it is often used in inter-computer networks for the transmission of format information and text layout [13]. As a rule, 7 or 8 bits are sufficient to store ASCII text in its most basic form. For example, TextEdit (Mac OS), SimpleText (Mac OS), Gedit (Unix, GNU/Linux), nano (Unix, GNU/Linux), Notepad (Windows), and ed (DOS) (Mac OS X). One may distinguish between two forms of "plain text" by looking at whether or not they have spaces. In the field of cryptography, the phrase "plaintext" refers to both the original text before encryption and the resulting encrypted text.

E. Jaro-Winkler Distance

Jaro-Winkler An method called "distance" is used to calculate how similar two strings are to one another. This method compares the two strings for similarities in order to find instances of repetition. A score of 0 indicates that no similarities were found between the strings, whereas a score of 1 indicates that they are the same. The three basic parts of this method are measuring the length of the string, finding strings of the same length, and enumerating the potential swaps [19]. To illustrate how [20] determines the separation (dj) between two strings (s1 and s2), consider the following:

$$d_j = \frac{1}{3} \times \left(\frac{m}{|s_1|} + \frac{m}{|s_2|} + \frac{m-t}{m} \right) \qquad (2)$$

Cover text is of length s1, stego text is of length s2, and the number of transpositions is t, where m is the total number of characters and s1 and s2 are the lengths of the respective text strings. In order to determine whether s1 and s2 are a match, check to see if their contents are same.

$$(3) \quad \left\lfloor \frac{\max |s_1|,|s_2|}{2} \right\rfloor - 1$$

With the following formula, you can find the Jaro-Winkler distance (dw) between two strings and get a good idea of how similar they are.

$$d_w = d_j + (lp(1 - d_j)) \qquad (4)$$

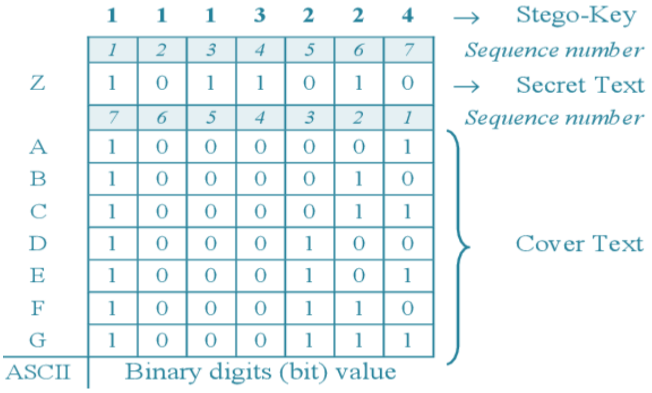
The Jaro-prefix Winkler's scale (p) is given a value of 0.1, and the prefix length (l) is equal to the same length as the character of the string being compared. This allows for a maximum inequality of 4 characters to be achieved, where s1 and s2 are the strings that are being compared to determine whether or not they are equal. The Jaro-prefix Winkler's scale is implemented with a value of 0.1 and a prefix length (l) that corresponds to the same number of characters as the string that is being evaluated for similarity [21.22].

3- Proposed Method: Bit Mapping on Ascii Characters

The binary digit (bit) of each secret text (ST) character is translated onto the binary digit (bit) of each cover text (CT) character, according to a technique that has been presented here. This allows the secret text (ST) to be concealed inside the cover text (CT) (CT). Within the context of this bit mapping, a binary digit stands in for every whitespace, punctuation mark, and symbol that is found in CT.[23].

A. Bit Mapping

When using the bit mapping technique of text steganography, the binary (bit) digit value is mapped onto each ASCII letter to generate a sequence number that may be used as a stego key. This approach is also known as the byte mapping method. In 7 bits, each character is represented by one of two values: either a ST or a CT value. The value in binary representation of the letter Z, for instance, is 1011010, which may be represented as ST, while the binary values of the letters ABCDEFG can be used as CT (1000001, 1000010, 1000011, 1000100, 1000101, 1000110, and 1000111). So, the sequence number is used as a steganography key, and the binary values of each ST character are encoded onto one of the identical binary values contained inside each CT character, reading from left to right. The complex process of mapping is seen in FIGURE 2 below.



ASCII	Binary digits (bit) value							
	1	1	1	3	2	2	4	→ Stego-Key
	1	*2*	*3*	*4*	*5*	*6*	*7*	*Sequence number*
Z	1	0	1	1	0	1	0	→ Secret Text
	7	*6*	*5*	*4*	*3*	*2*	*1*	*Sequence number*
A	1	0	0	0	0	0	1	Cover Text
B	1	0	0	0	0	1	0	
C	1	0	0	0	0	1	1	
D	1	0	0	0	1	0	0	
E	1	0	0	0	1	0	1	
F	1	0	0	0	1	1	0	
G	1	0	0	0	1	1	1	

Figure 2 Bit mapping

B. Method

The following is a description of the bit mapping technique of text steganography: If we have an alphabet, A = {A,B,C,...,Y,Z,} a set of integers, Z26, a collection of secret texts, M = a1,a2,...,am with ai€A, and a set of one-time password keys, Ki=k1,k2,...,km with ki€A, then the function of Asc(x) is the function that converts xA into the final seven digits of the ASCII code [24,25]. You may decrypt message E by:

$$E = f^{-1}([f(M) + f(K)]\ mod\ 26) \qquad (5)$$

Following this, we may derive the binary representation E by:

$$E_b = Asc(e), \forall e \in E \qquad (6)$$

with |Eb| = 7m. For example, cover-text used C=c1, c2 , cn

where n is greater than or equal to 7m. One may convert C into its binary version by doing:

$$C_b = Asc(c), \forall c \in C \qquad (7)$$

Following these criteria, the binary digits Eb are embedded inside the binary digits Cb. Reading Eb proceeds from left to right, whereas reading Cb begins with the initial character (the first seven numbers) and proceeds counterclockwise. For instance, if i = 1,2,.7m and j = 1,2,.7n, then eiEb and cjCb, respectively.

The stego key is then derived via an embedding procedure W=w1,w2,...,w7n with rules wk=j if ei=cj, yielding a stego-text S that is visually indistinguishable from the cover-text.

If a message receiver, M, receives a stego message (S) together with a stego key (W) and a one-time password (OTP) key (K), M may decipher M using the following steps. First, we use it to transform the steganography text S into the binary representation Sb.

$$S_b = Asc(s), \forall c \in S \qquad (8)$$

Coded Document By employing stego key W, we can decrypt Sb and get Eb. To demonstrate, we may get the index digit from wk on Sb (k=1,2,...,7n) and use it to extract the corresponding character (binary digit 7) in Eb. Following this, Eb is changed into E using

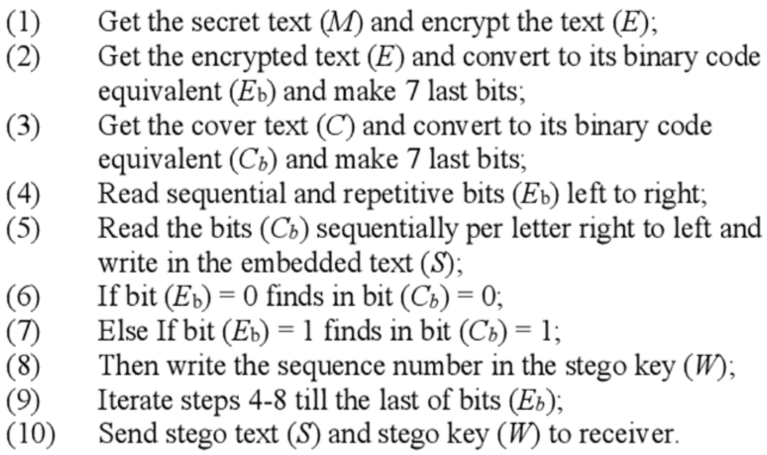
$$E = Asc^{-1}(e), \forall e \in E_b \qquad (9)$$

subsequently E is converted into M with

$$M = f^{-1}([f(E) - f(K)]\ mod\ 26) \qquad (10)$$

That means we can now read the hidden message that was concealed inside the cover text.

Figure 4 depicts the scheme of the bit mapping technique applied to the ASCII letters, while Figures 2 and 3 display the steps of the embedding process and the extraction methodology, respectively.



```
(1)   Get the secret text ($M$) and encrypt the text ($E$);
(2)   Get the encrypted text ($E$) and convert to its binary code
      equivalent ($E_b$) and make 7 last bits;
(3)   Get the cover text ($C$) and convert to its binary code
      equivalent ($C_b$) and make 7 last bits;
(4)   Read sequential and repetitive bits ($E_b$) left to right;
(5)   Read the bits ($C_b$) sequentially per letter right to left and
      write in the embedded text ($S$);
(6)   If bit ($E_b$) = 0 finds in bit ($C_b$) = 0;
(7)   Else If bit ($E_b$) = 1 finds in bit ($C_b$) = 1;
(8)   Then write the sequence number in the stego key ($W$);
(9)   Iterate steps 4-8 till the last of bits ($E_b$);
(10)  Send stego text ($S$) and stego key ($W$) to receiver.
```

Figure 3 Pseudocode for embedding algorithm

```
(1)   Get the stego text ($S$) and convert to its binary code
      equivalent ($S_b$);
(2)   Get the stego key ($W$) and read sequentially;
(3)   Read consecutive bits ($S_b$) per letter;
(4)   If key ($W$) = 1 then take the first bit ($S_b$);
(5)   If key ($W$) = 2 then take the second bit ($S_b$);
(6)   If key ($W$) = 3 then take the third bit ($S_b$);
(7)   If key ($W$) = 4 then take the fourth bit ($S_b$);
(8)   If key ($W$) = 5 then take the fifth bit ($S_b$);
(9)   If key ($W$) = 6 then take the sixth bit ($S_b$);
(10)  Else If key ($W$) = 7 then take the seventh bit ($S_b$);
(11)  Then write in the file ($E_b$);
(12)  Repeat steps 3-11 till the last of bits ($S_b$);
(13)  Convert every 7 bits the file ($E_b$) to its ASCII binary
      code equivalent and write to encrypted text ($E$);
(14)  Obtained encrypted text ($E$);
(15)  Decrypt ($E$) to secret message (M).
```

Figure 4 Pseudocode for the extracting

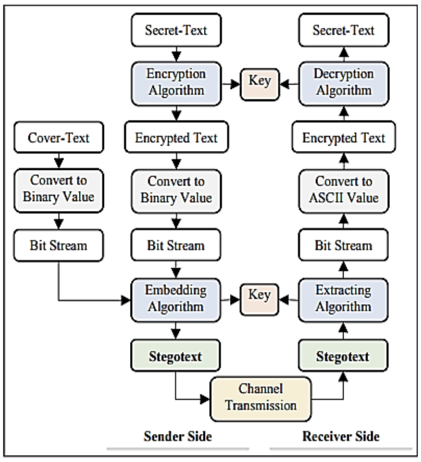


Figure 5 proposed method processing

C. Algorithm

Example

An OPT encryption algorithm was used to encrypt the ST text before it was turned into a bit stream and used further. The CT text, unlike the ST text, was instantly transformed into a digital format. After that, the stego text and key were embedded. This procedure yielded the following results:

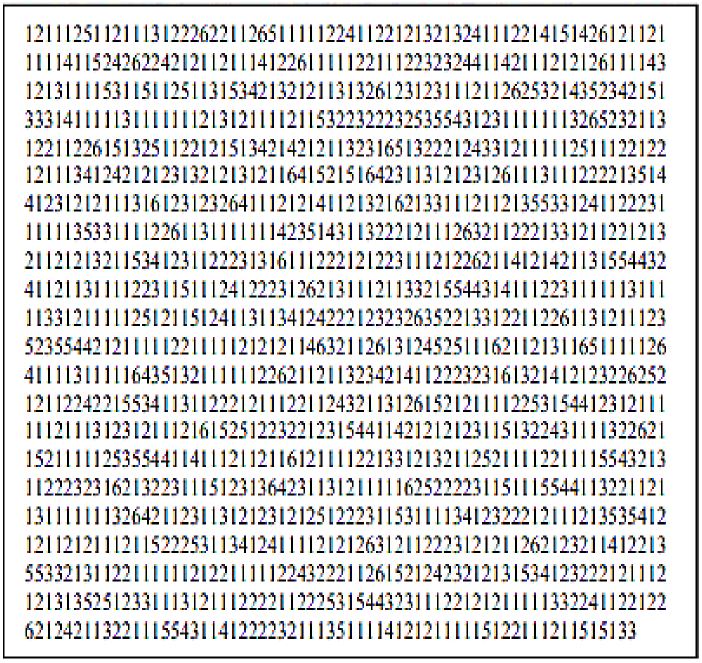
- Secret Text

VisiKementerianAgamaRepublikIndonesiaTerwujudnyaMasyarakatIndonesiaYangTaatBeragamaRukunCerdasDanSejahteraLahirBatinDalamRangkaMewujudkanIndonesiaYangBerdaulatMandiriDanBerkepribadianBerlandaskanGotongRoyong

- OTP Key:

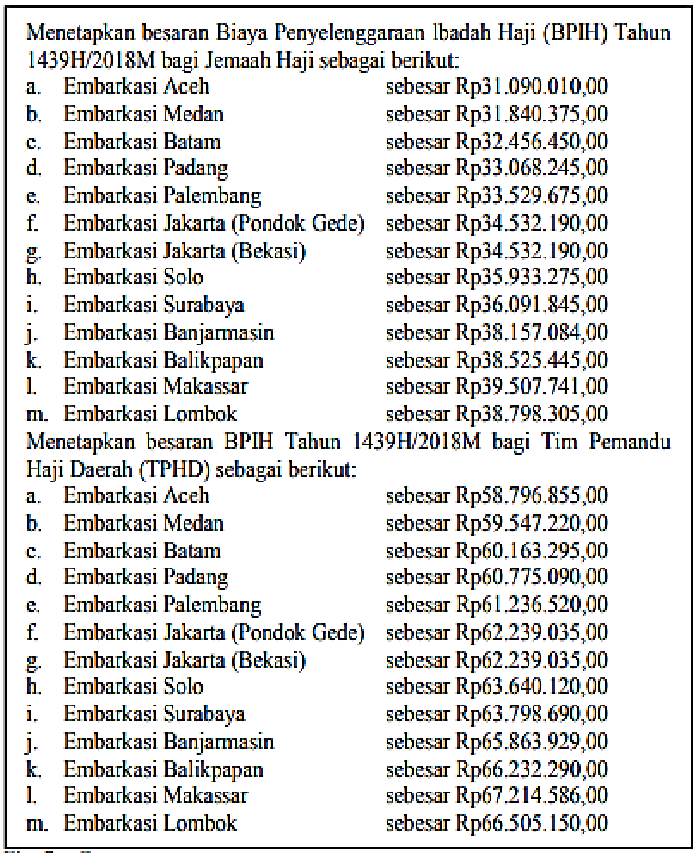
Njvfuihaurfjannfvueheqfjnvifhvaerjfrfurufgrehfiaghufgjfnghpigjrezmnxbcvvhfggdfjdkslapoqiqwuieurytttgfghfdjdkzmxnxbcvvgfhdjskslapqoqiwiuewuerutyttqpplkjdhdiedutethvhvbcgcgpoiiuyuttwqwadsfdgfhjklklmnbvcxcvjnjd

The Stego key is in Figure 5, the cover text is in Figure 6, and the stego text is in Figure 7.



121112511211131222622112651111122411221213213241112214151426121121
111141152426224212112111412261111122111223232441142111212126111143
121311115311511251131534213212113132612312311121126253214352342151
333141111131111111213121111211532232223253554312311111113265232113
122112261513251122121513421421211323165132221243312111112511122122
121113412421212313212131211641521516423113121231261113111222213514
412312121113161231232641112121411213216213311121121355331241122231
11111353311122611311111114235143113222121112632112221331211221213
211212132115341231122231316111222121223111212262114121421131554432
411211311112231151112412223126213111211332155443141112231111113111
113312111112512115124113113412422212323263522133122112261131211123
523554421211111221111121212114632112613124525111621121311651111126
411113111116435132111111226211211323421411222323161321412123226252
121122422155341131122212111221124321131261521211112253154412312111
111211131231211121615251223221231544114212121231151322431111322621
152111112535544114111211211612111122133121321125211112211115543213
112223231621322311151231364231131211111625222231151115544113221121
1311111111326421123113121231212512223115311113412322212111213535412
121121211121152225311341241111212126312112223121211262123211412213
553321311221111112122111112243222112615212423212131534123222121112
121313525123311131211122221122253154432311122121211111332241122122
621242113221115543114122223211135111141212111115122111211515133

Figure 4 Stagey



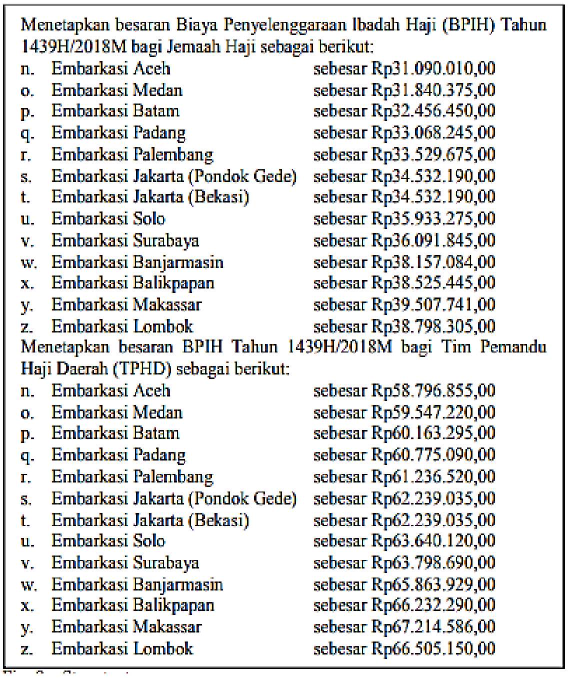
Menetapkan besaran Biaya Penyelenggaraan Ibadah Haji (BPIH) Tahun 1439H/2018M bagi Jemaah Haji sebagai berikut:

a.	Embarkasi Aceh	sebesar Rp31.090.010,00
b.	Embarkasi Medan	sebesar Rp31.840.375,00
c.	Embarkasi Batam	sebesar Rp32.456.450,00
d.	Embarkasi Padang	sebesar Rp33.068.245,00
e.	Embarkasi Palembang	sebesar Rp33.529.675,00
f.	Embarkasi Jakarta (Pondok Gede)	sebesar Rp34.532.190,00
g.	Embarkasi Jakarta (Bekasi)	sebesar Rp34.532.190,00
h.	Embarkasi Solo	sebesar Rp35.933.275,00
i.	Embarkasi Surabaya	sebesar Rp36.091.845,00
j.	Embarkasi Banjarmasin	sebesar Rp38.157.084,00
k.	Embarkasi Balikpapan	sebesar Rp38.525.445,00
l.	Embarkasi Makassar	sebesar Rp39.507.741,00
m.	Embarkasi Lombok	sebesar Rp38.798.305,00

Menetapkan besaran BPIH Tahun 1439H/2018M bagi Tim Pemandu Haji Daerah (TPHD) sebagai berikut:

a.	Embarkasi Aceh	sebesar Rp58.796.855,00
b.	Embarkasi Medan	sebesar Rp59.547.220,00
c.	Embarkasi Batam	sebesar Rp60.163.295,00
d.	Embarkasi Padang	sebesar Rp60.775.090,00
e.	Embarkasi Palembang	sebesar Rp61.236.520,00
f.	Embarkasi Jakarta (Pondok Gede)	sebesar Rp62.239.035,00
g.	Embarkasi Jakarta (Bekasi)	sebesar Rp62.239.035,00
h.	Embarkasi Solo	sebesar Rp63.640.120,00
i.	Embarkasi Surabaya	sebesar Rp63.798.690,00
j.	Embarkasi Banjarmasin	sebesar Rp65.863.929,00
k.	Embarkasi Balikpapan	sebesar Rp66.232.290,00
l.	Embarkasi Makassar	sebesar Rp67.214.586,00
m.	Embarkasi Lombok	sebesar Rp66.505.150,00

Figure 5 Cover-text



Menetapkan besaran Biaya Penyelenggaraan Ibadah Haji (BPIH) Tahun 1439H/2018M bagi Jemaah Haji sebagai berikut:

n.	Embarkasi Aceh	sebesar Rp31.090.010,00
o.	Embarkasi Medan	sebesar Rp31.840.375,00
p.	Embarkasi Batam	sebesar Rp32.456.450,00
q.	Embarkasi Padang	sebesar Rp33.068.245,00
r.	Embarkasi Palembang	sebesar Rp33.529.675,00
s.	Embarkasi Jakarta (Pondok Gede)	sebesar Rp34.532.190,00
t.	Embarkasi Jakarta (Bekasi)	sebesar Rp34.532.190,00
u.	Embarkasi Solo	sebesar Rp35.933.275,00
v.	Embarkasi Surabaya	sebesar Rp36.091.845,00
w.	Embarkasi Banjarmasin	sebesar Rp38.157.084,00
x.	Embarkasi Balikpapan	sebesar Rp38.525.445,00
y.	Embarkasi Makassar	sebesar Rp39.507.741,00
z.	Embarkasi Lombok	sebesar Rp38.798.305,00

Menetapkan besaran BPIH Tahun 1439H/2018M bagi Tim Pemandu Haji Daerah (TPHD) sebagai berikut:

n.	Embarkasi Aceh	sebesar Rp58.796.855,00
o.	Embarkasi Medan	sebesar Rp59.547.220,00
p.	Embarkasi Batam	sebesar Rp60.163.295,00
q.	Embarkasi Padang	sebesar Rp60.775.090,00
r.	Embarkasi Palembang	sebesar Rp61.236.520,00
s.	Embarkasi Jakarta (Pondok Gede)	sebesar Rp62.239.035,00
t.	Embarkasi Jakarta (Bekasi)	sebesar Rp62.239.035,00
u.	Embarkasi Solo	sebesar Rp63.640.120,00
v.	Embarkasi Surabaya	sebesar Rp63.798.690,00
w.	Embarkasi Banjarmasin	sebesar Rp65.863.929,00
x.	Embarkasi Balikpapan	sebesar Rp66.232.290,00
y.	Embarkasi Makassar	sebesar Rp67.214.586,00
z.	Embarkasi Lombok	sebesar Rp66.505.150,00

Figure 6 Steg text

4. *Results and Discussion :*

What follows is a sample of the study's findings and the discussion that follows their application of the bit mapping approach.

A. Experimental

In order to assess how much ST can be handled by CT and whether or not embedding it into plain text would modify the plain text itself, an experiment was carried out on an information medium that was based on plain text.After a stego key was created, ST could be reconstructed from the resulting stego text.

The testing made use of article texts sourced from a broad variety of online mass media sources, ecommerce websites (used in the product description part), major social media, and SMS. Table I displays the findings obtained from the experiment.

Table 1 COMPARISON COVER TEXT CAPACITY

A piece of plain text	Size as cover text		Secret text in bytes	Capacity %	Jaro score
	Bytes	Max / avg			
Detik.com	2241	Average	320	14.279	1
Kompas.com	2428	Average	346	14.250	1
Bukalapak	1094	Average	156	14.265	1
Tokopedia	1101	Average	157	14.260	1
SMS	160	Maximum	22	13.750	1
Twitter	280	Maximum	40	14.286	1
Facebook	63206	Maximum	9029	14.285	1
WhatsApp	65536	Maximum	9362	14.285	1

The average size of online news and ecommerce content was calculated by randomly sampling 100 articles from each area, Each of which had a unique assortment of ASCII characters and lengths of paragraphs.

Since the character counts (CT) of social media posts and SMS messages are capped by default, the latter is used to signify that certain procedures are still viable even on medium with character counts (CT) that are much lower, such as SMS.

B- Capacity

By calculating the capacity ratio and solving equation (11), which states that 1 bit of ST can be stored in 7 bits of CT and 1 ASCII character can hold 1 bit of ST, we can determine that 1 bit of ST can be stored in 7 bits of CT, we may determine that CT can store ST. There is a 14.2% embedding capacity.

$$Capacity\ ratio = \frac{bits\ of\ secret\ text}{bits\ of\ cover\ text} \times 100\%$$

C. Similarity measure

Using Jaro-Winkler Distance, where a score of 0 indicates no similarity and a score of 1 indicates exactly the same, the generated stego-text was compared to CT once the embedding process was complete. Using equation (2), we can see that the similarity between the two strings is 1, indicating that they are identical.

D. Discussion

Steganography's primary goal is to covertly transmit sensitive information through a chosen medium. The best methods of concealment are those that are not immediately obvious.

The best strategy is the one that results in a stego-file that is almost identical to its cover file. The stego-text created by this bit-mapping technique is identical to the CT. Using the ASCII character's binary value, an undetectable stego-text may be created that seems to be a standard string of letters but really contains a hidden message.

This is a two-key technique for securing ST; the OTP key encrypts ST before it is implanted in CT, and the stego-key keeps it hidden after it is within CT. The symmetric encryption with a statistically random key makes this approach very secure.

With a capacity ratio of 14.2% and a 7-bit CT's ability to contain 1 bit ST, the results are a greater embedding capacity. Since bit mapping can be used on any plain text, this text steganography technique could be used for a wide range of digital purposes, such as sending secret messages and keeping plain text information safe.

Due to the many limitations of character recognition, especially on non-printable characters such as Tab, CR and LF, the hidden ST in CT may also be defective and disappear when printing and scanning with OCR, but the hidden ST in CT will not be defective and disappear when copied and pasted On a wide variety of digital platforms, such as Microsoft Word, Google Docs and other similar applications.

Since the ST bit is used to insert each CT character, this technique may be used to prevent tampering with plain text by identifying when the CT character sequence has been altered by insertion, deletion, or substitution. This method makes plain text fragile, which makes sure that plain text is real and can be trusted.

5. Conclusions

This bit mapping technique efficiently converts secret-text (ST) into cover-text (CT) with a capacity ratio of 14.2% by converting characters into bit streams. So, ST bits may be masked by any letter in the ASCII alphabet. Jaro-Winkler Distance Measurement reveals that the bit mapping method's stego-text is qualitatively equivalent to CT [1].

Aside from the embedded ST, the stego-text is indistinguishable from the original CT. We relied on accessible writing throughout the Internet, from news stories and social media postings to e-commerce site descriptions, to conduct our analysis.

We also relied on the unabridged version of the Decree of the Minister of Religious Affairs.

The experiments showed that there isn't much difference between the two, and the bit mapping method works just as well with either one as a cover text.

This research could lead to a deeper look at character encodings other than the standard ASCII.

References

[1] A. Naharuddin, A. D. Wibawa and S. Sumpeno, "A High Capacity and Imperceptible Text Steganography Using Binary Digit Mapping on ASCII Characters," 2018 International Seminar on Intelligent Technology and Its Applications (ISITIA), Bali, Indonesia, 2018, pp. 287-292, doi: 10.1109/ISITIA.2018.8711087

[2] J. l.Azzam, S. N., & Al-Garni, F. A. (2023). The use of binary digit mapping on ASCII characters to create a high-capacity, undetectable text steganography. Journal of Advanced Sciences and Engineering Technologies, 5(2), 49–60. https://doi.org/10.32441/jaset.05.02.05

[3] Naharuddin, A., Wibawa, A. D., & Sumpeno, S. (2018). A High Capacity and Imperceptible Text Steganography Using Binary Digit Mapping on ASCII Characters. 2018 International Seminar on Intelligent Technology and Its Applications (ISITIA). https://doi.org/10.1109/isitia.2018.8711087

[4] P. Johri, A. Mishra, and S. Das, “Survey on steganography methods (text, image, audio, video, protocol and network steganography),” 2016 3rd Int. Conf. Comput. Sustain. Glob. Dev., pp. 2906–2909, 2016.

[5] S. Sharma, A. Gupta, M. C. Trivedi, and V. K. Yadav, “Analysis of different text steganography techniques: A survey,” Proc. - 2016 2nd Int. Conf. Comput. Intell. Commun. Technol. CICT 2016, pp. 130–133, 2016.

[6] Elshoush, H. T., Mahmoud, M. M., & Altigani, A. (2021). A new high capacity and secure image realization steganography based on ASCII code matching. Multimedia Tools and Applications,

81(4), 5191–5237. https://doi.org/10.1007/s11042-021-11741-y

[7] Al Smadi, T. A. (2013). Design and Implementation of Double Base Integer Encoder of Term Metrical to Direct Binary Code Application. Journal of Signal and Information Processing, 4(4), 370-374.

[8] Al Smadi, T. A., & Maitah, M. (2014). An efficiency and algorithm detection for stenography in digital symbols. International Journal of Computer Network and Information Security, 6(1), 34-38.

[9] T. Acharjee, A. Konwar, R. K. Ram, R. Sharma, and D. Goswami, "XORSTEG: A New Model of Text Steganography," 2016.

[10] Ka. Kumar ProfSuresh Pabboju, "A Comparative Result Analysis of Text Based Steganographic Approaches," IOSR J. Comput. Eng. Ver. VII, vol. 17, no. 3, pp. 2278–661, 2015.

[11] A Coverless Text Steganography by Encoding the Chinese Characters' Component Structures. (2021). International Journal of Digital Crime and Forensics, 13(6), 0–0. https://doi.org/10.4018/ijdcf.302135

[12] Smadi, T. A., Al-Khawaldeh, I., & Smadi, K. A. (2017). Three-Dimensional Scenes Restore Using Digital Image. Journal of Signal and Information Processing, 08(01), 1–8. https://doi.org/10.4236/jsip.2017.81001

[13] P. Christensson, "Plain Text Definition," TechTerms., 2010. [Online]. Available: https://techterms.com/definition/plaintext. [Accessed: 14- May-2017].

[14] Al-Smadi, T. A., & Al-Wahshat, H. (2011). System identification of the logical object and logical acupuncture. International Journal of Physical Sciences, 6(15), 3771-3777.

[15] B. Leonardo and S. Hansun, "Text documents plagiarism detection using Rabin-Karp and Jaro-Winkler distance algorithms," Indones. J. Electr. Eng. Comput. Sci., vol. 5, no. 2, pp. 462–471, 2017.

[16] Al-Smadi, T. A., & Ibrahim, Y. K. (2007). Design of Speed Independent Ripple Carry Adder. Journal of Applied Sciences, 7(6), 848-854.

[17] Ghaidan, K. A., Al Smadi, T. A., Aljumailly, T. A., & Al-Taweel, F. M. (2011). Development of a new approach for transmitting Digital message on a Frequency Limited Communication Channel transmission. ournal of Advanced Computer Science and Technology Research, 1, 52-62.

[18] Naharuddin, Alfin, Adhi Dharma Wibawa, and Surya Sumpeno. "A high capacity and imperceptible text steganography using binary digit mapping on ASCII characters." 2018 International Seminar on Intelligent Technology and Its Applications (ISITIA). IEEE, 2018.

[19] Elshoush, H. T., Mahmoud, M. M., & Altigani, A. (2022). A new high capacity and secure image realization steganography based on ASCII code matching. Multimedia Tools and Applications, 1-47.

[20] Du, G. E. J. (2016). Ensuring the Security of Text-Based Information Transmission by Utilizing Invisible ASCII Characters. International Journal of Simulation: Systems, Science &amp; Technology. https://doi.org/10.5013/ijssst.a.17.45.34

[21] Vo, P.-H., Nguyen, T.-S., Huynh, V.-T., & Do, T.-N. (2020). A high-capacity invertible steganography method for stereo image. Digital Media Steganography, 99–122. https://doi.org/10.1016/b978-0-12-819438-6.00014-1

[22] Vo, P.-H., Nguyen, T.-S., Huynh, V.-T., & Do, T.-N. (2020). A high-capacity invertible steganography method for stereo image. Digital Media Steganography, 99–122. https://doi.org/10.1016/b978-0-12-819438-6.00014-1

[23] A. Al Smadi, T., & Maitah, M. (2013). An Efficiency and Algorithm Detection for Stenography in Digital Symbols. International Journal of Computer Network and Information Security, 6(1), 34–38. https://doi.org/10.5815/ijcnis.2014.01.05

[24] Naharuddin, A., Wibawa, A. D., & Sumpeno, S. (2018, August). A high capacity and imperceptible text steganography using binary digit mapping on ASCII characters. In 2018 International Seminar on Intelligent Technology and Its Applications (ISITIA) (pp. 287-292). IEEE.

[25] Bashir, H. M., Li, Q., & Hou, J. (2020, November). A high capacity text steganography utilizing unicode zero-width characters. In 2020 International Conferences on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics) (pp. 668-675). IEEE.