More Secured Text Transmission with Dual Phase Message Morphing Algorithm

Article · April 2012

CITATIONS
0

READS
4

3 authors, including:

Dr. Meka James Stephen
WELLFARE INSTITUTE OF SCIENCE TECHNOLOGY AND MANAGEMENT (WISTM)
22 PUBLICATIONS **52** CITATIONS

SEE PROFILE

More Secured Text Transmission with Dual Phase Message Morphing Algorithm

M. James Stephen[1], P.V.G.D. Prasad Reddy[2], Ch. Demudu Naidu[3], Sampangi Sonali[4], and Ch. Heymaraju[5]

[1] Department of I.T, ANITS, Visakhapatnam, India
jamesstephenm@yahoo.com
[2] Dept. of CS & SE, Andhra University, Visakhapatnam, India
prasadreddy.vizag@gmail.com
[3] Department of I.T, ANITS, Visakhapatnam, India
naidu.061@gmail.com
[4] Dept. of I.T, ANITS
[5] Dept.of I.T, GIT, G.U

Abstract. In this global village, where with the advancements in the field of communications, ensuring security to the data being transmitted has become very vital. These safety and security issues contributed to the outgrowth of secret communication. While encryption simply encodes the data making it difficult for the layman to understand, Steganography deals with hiding data within another data making them unaware of it. Both concepts ensure data security but in different forms. These two approaches when unified provide much better security to the data than that provided by either of encryption or Steganography.

This paper proposes such a new unified approach for Secured Text Transmission using Dual Phase Message Morphing (DPMMA) algorithm, which encrypts and conceals data in two consecutive stages to provide better security to the data. This algorithm is a simple unified approach of encryption and Steganography which employs two newly proposed techniques for encryption and Steganography to provide better security for the data.

As the name suggests it works in two consecutive phases. In the first phase encryption is performed and in the second phase the encrypted message is concealed within another text. The result of these two phases produces a morphed text which does not resemble the original message.

Keywords: E-Message, Message, Cover Message, Sequence, Label, Encryption, Text Steganography.

1 Introduction

The advent of internet has benefited the human life up to a greatest extent by connecting the entire globe on a single medium. E-mails, chat rooms, social networking sites, data sharing etc all these features catalyzed the widespread of internet. Many attackers try to intrude into the network and capture the data being transmitted. Thus providing security to data being transmitted has turned out to be a

S.C. Satapathy et al. (Eds.): Proceedings of the InConINDIA 2012, AISC 132, pp. 845–852.
springerlink.com

crucial task during communications. The best opted methods for secret communications are Data Encryption and Steganography. Steganography is the art and science of hidden writing. While an encryption program protects your message from being read by those not in possessions of the key, sometimes you wish to obscure the very fact you're sending an encrypted message at all.

The comparison of different techniques for communicating in secret can be found in [1]. The unified approach of the above two techniques will provide a better security. Several existing Steganography methods can encrypt data before hiding it in the chosen medium [2]. But all these methods use image as their cover data, where as the present method employees text as the cover data.

2 Why Text Steganography?

Text Steganography deals with hiding information in simple text data but not digital data. It is relatively easy method to hide data and makes the process of breaking the hidden message difficult unless technique used is known. The comparison of various forms of Steganography is as shown in the Table 2 [3].

Table 1. Comparision of Various Forms of Steganography

Steganography Techniques	Medium	Embedding Technique
Steganography Using Text	Text files	To embed information we need to simply alter the text to a suitable form.
Steganography Using Audio	MP3 files	Encode data as a binary sequence which sounds like noise
Steganography Using Image	Image files	It works by altering the bit configurations or wavelets.
Steganography Using Video	Video files	A combination of sound and image techniques can be used.

From the above table we can see that text Steganography is the simplest form of Steganography. The best thing about this text Steganography is that the hidden information cannot be extracted unless the mode of embedding is known.

There are number of techniques existing techniques for text Steganography [4] like 'Steganography of Information in Random Character and Word Sequences' , 'Steganography of Information in Specific Characters in Words' ,'Creating Spam Texts', 'Line Shifting', 'Word Shifting' , 'Syntactic Methods' , 'Semantic Methods', 'Feature Coding', 'Abbreviation', 'Open Spaces', 'Persian/Arabic Text Steganography'

Rather than placing the alphabet in particular positions, it would be a better idea to place the characters at desired random positions making it difficult for the attacker to break the hidden message. Hence the message is secured. What if the message is

encrypted before placing it in random positions? Then the message becomes doubly secured. This particular point works as the basic principle for our suggested DPMMA algorithm.

3 Proposed Work

Here we present a new method for text Steganography using DPMMA (Dual Phase Message Morphing Algorithm).This algorithm is a simple unified approach which employs two newly proposed techniques for encryption and Steganography to provide better security for the data.

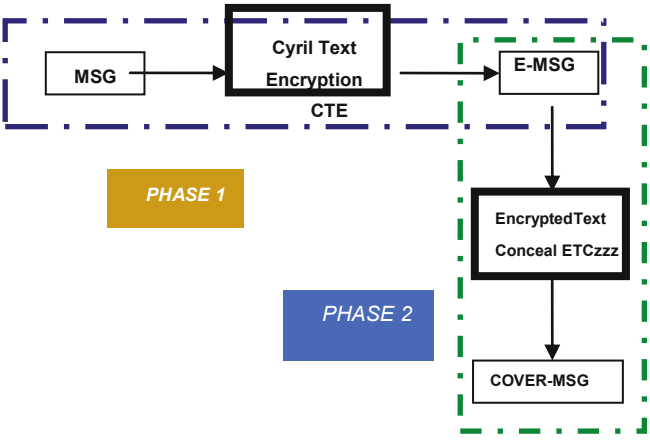

Fig. 1. Structure of DPMMA

As the name suggests it works in two consecutive phases. In the first phase encryption is performed and in the second phase the encrypted message is concealed within another text. The result of these two phases produces a morphed text which doesn't resemble the original message.

3.1 First Phase of DPMMA

The first stage performs encryption. The original text (Message) is encrypted to produce the cipher text (E -Message). Encryption is the method of changing the text from its original form to another form that cannot be easily understood. Substitution techniques are the simplest encryption techniques. In this algorithm a new substitution technique named CTE (Cyril Text Encryption) is introduced.

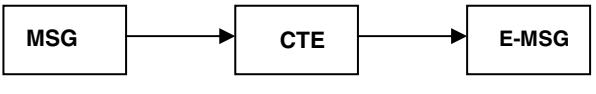

Fig. 2. First phase of DPMMA

CTE is a substitution technique based on a table called Cyril Tab. This table is generated based on a Russian font called Cyrillic font. The character in the original text (Message) get substituted by another character depending on the table thus producing the encrypted text (E-Message).

Table 2. CYRIL TAB

a	b	c	d	E	f	g	h	i	J	k	l	m
f	u	c	t	Y	a	n	p	i	O	z	d	b
n	o	p	q	R	s	t	u	v	W	x	y	z
m	w	x	q	K	g	e	s	v	J	l	h	r

Special characters and numeric data if used in message are not encrypted but remain the same in e-message also. Based on the above table the substitution is performed and the E-Message is now passed to next phase to create Cover message.

Algorithm for Encryption Using Cyril Tab.

```
a : message
e : e-message
start  cte ( a )
for each character a[i] do
encrypt a[i] using cyril tab
end for
return e
end  cte
```

Fig. 3. Pseudo code of CTE technique

3.2 Second Phase

Johnson and Katzenbeisser grouped steganographic techniques into six categories depending on how the algorithm encodes information in the cover object. They are: substitution systems, transform domain techniques, spread spectrum techniques, statistical methods, distortion techniques, and cover generation methods[5]. Cover generation techniques are most unique of these six types. A cover generation method actually creates a cover for the sole purpose of hiding information.

This phase deals with concealing the E-Message generated from the first phase. To conceal the E-Message we generate another text called the Text Passage using ETC (Encrypted Text Conceal) technique which is a new cover generation steganography technique. This technique involves two important features, one is Text Passage and the other is Label which will be seen later.

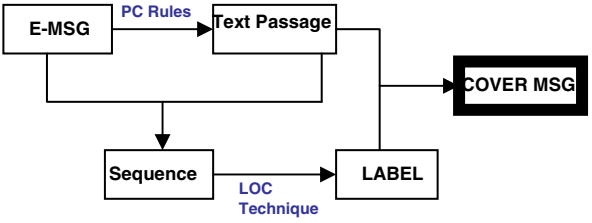

Fig. 4. Second phase of DPMMA

To generate the Text Passage from the E-Message certain rules are defined called as PC (Passage Creation) Rules.

Rule 1: Each alphabet in the E-Message word should correspond to each word in the Text Passage.
Rule 2: Each word in the E-Message should correspond to each sentence in the Text Passage.
Rule 3: Every alphabet in the E- Message can be placed in any desired position (between 1 to 9) in the words in the Text Passage.
Rule 4: Every sentence in the text passage should end with Full stop and unwanted spaces should be avoided.

Algorithm for labeling the cover message

```
e : e-message
p : text passage
s: sequence
start etc( e , p)
w=split p to array of strings;
for each e[i]  do
    s[i]=index of e[i] in s[i]
end for
label=loctechnique( s[i] )
add  label to p
end etc
```

Fig. 5. ETC technique

Based on the above rules the text passage is generated. It is made clear that every word in the passage contributes to the hidden message. At this point it is understood that hidden message is embedded within the passage at random positions (between 1 to 9) as desired by the user. The positions at which the data in the E-Message is placed, when combined together produce a Sequence. This sequence is actually generated from the E-Message and the Text Passage. Sequence plays the key role in extracting the hidden data from the passage. Hence this sequence also needs to be secured.

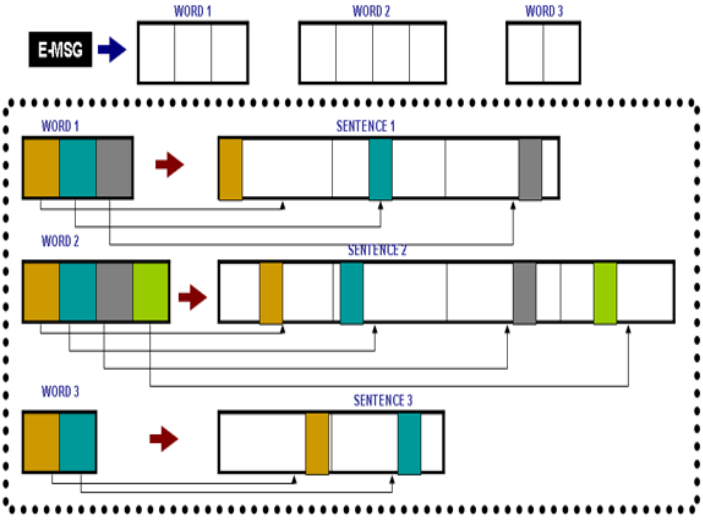

Fig. 6. PC Rules

(Long Octal conversions) technique. This technique is an encryption technique used for the encryption of numeric data. In this technique the sequence is initially converted to a set of long numbers which are then converted to their corresponding octal strings and finally then concatenated to produce a string called Label.

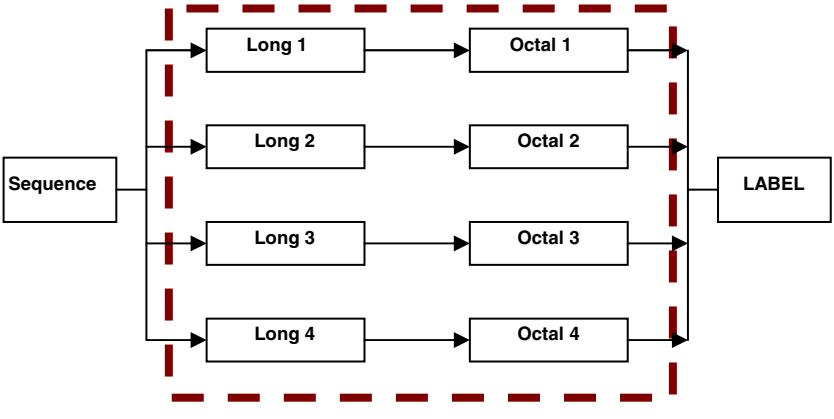

Fig. 7. LOC Technique

This Label is then appended at the end of the text passage to create the cover data (Cover Message) which can then be sent to the required person. This Label concept is analogous to digital signature but method of computation is totally different. This completes the generation of cover message using ETC technique. Now this cover message which is the required cover data can be used for secret communications.

Algorithm for LOC technique

```
start loctechnique(  s )
initialize k=i+19 and label,octal as null;
for each i less than equal to s.length do
l[j]=convert s[i] to s[k] as long and adjust I,j,k
end for
for each l[i] do
    o[i]=octal string of l[i]
    octal=add octal and o[i]
end for
return octal
end  loctechnique
```

Fig. 8. LOC technique

Extracting the data from the cover message can be performed in the reverse process of the employed techniques. The label is first extracted from the cover message and then it is decrypted to produce the sequence. The sequence is then used to extract characters from the passage. The extracted data when decrypted produce the original hidden message.

Algorithm for Extraction of Data from cover message

```
start extract ( p, s )
l=label,o= octal strings
for each i less than equal to o.length do
l[i]=convert o[i] to its equivalent long
end for
for each j less than l.length
s[k]=convert l[i] to set of integers
end for
w=split p to array of strings;
for each s[i]  do
    a[i]=character at s[i] in w[i]
end for
decrypt a[i] using cyril tab
 end extract
```

Fig. 9. Pseudo code of extraction

4 Results

As discussed the algorithm works in two consecutive stages. The first stage performs encryption to produce encrypted message using CTE technique. The number of characters in the message and e-message remain same. So no data loss can occur.

The second stage uses cover generation Steganography technique called ETC to generate the cover message along with label. Each character in e-message should correspond to each word in passage, and label is generated only when the number of characters and words match, hence data cannot be lost.

Our experiments proved that this proposed system is 99 % successful in providing better .99.5 % attacks to break the message completely resulted in failure. Since the system was developed entirely using Java, it works on any platform that supports java.

5 Conclusion

Security and privacy issues enhanced the growth of secret communications. Many Steganography techniques evolved to transmit hidden messages and at the same time lot of analysis works were carried out to detect the presence of hidden messages which raised the need for more robust Steganography approaches.

This paper proposes a robust and secure method of Steganography using DPMMA which encrypts and then hides the data. The cover data can be transmitted over mails, chats, text files , contents of web pages etc, where mostly text is used.

6 Future Implementation

This is a flexible user-friendly application developed to provide robust and secure transmissions of hidden data. We look forward to add a word suggestion algorithm to make it much more user friendly. Because of it flexibility and efficiency it can be added as a feature in mail systems or as an add-on on the web browsers.

References

[1] Cummins, J., Diskin, P., Lau, S., Parlett, R.: Steganography and digital watermarking. School of Computer Science, The University of Birmingham (2003)
[2] Artz, D.: Digital Steganography: hiding data within data. IEEE Internet Computing (May-June, 2001)
[3] Channalli, S., et al.: International Journal on Computer Science and Engineering 1(3) (2009)
[4] Shirali Shahreza, M.: Text steganography in sms. In: International Conference on Convergence Information Technology (July 2007)
[5] Bennett, K.: Linguistic Steganography: Survey, Analysis, and Robustness Concerns for Hiding Information in Text, Purdue University, CERIAS Tech. Report 2004-13 (2004)