Information Hiding in Email Services Based on Confused Document Encrypting Schemes

Author:
Wei-Shyun Pan
Po-Kang Chen
Quincy Wu
• Introduction
• Related works
  – Information hiding
• System model
  – Concept
  – CDES module
  – Compression module
  – Encryption module
  – Image hiding module
• Implementation
  – Environment
  – Experiment
• Conclusion & Future work
Introduction

• The “Personal Privacy” becomes a popular section in information security over Internet.

• **Information Hiding** (Steganography) and **Cryptography** accomplish secret communication between you and me.
The Eavesdropper  (Nicolaes Maes, 1657)
Motivation

• Provide a secret communication service for Email over Internet and demonstrate how Email services can be protected in my system.

• CDES (Confused Document Encrypting Scheme) is a technique for data hiding, which sends a meaningful message to deceive the eavesdropper and increase the security.

• Add the image hiding technique.
Related works
• Use any media to hide secret information.
• the hacker cannot sense something when he intercepts the media, because it is common behavior.

Secret information
• Confused Document Encrypting Scheme (CDES), Lin & Lee, 1998 [1]

• Elements of CDES
  – Cheating texts
  – Plaintext
  – Character position table (CPT)
  – Plaintext Index file (PIF)
  – Key
Transmitting many cheating text files

Plaintext index file (PIF)

From: weishin.pan@gmail.com
To: magicpanx@gmail.com
Subject: Hello!

Confused Document Encrypting Scheme
Attachments:
(1) ID-0005.txt
(2) ID-0019.txt
(3) ID-5597.txt
(4) ID-2468.txt
(5) ID-9870.txt
(5) Encrypted plaintext index file + (Encrypted-ID 0019)
Sender

1. Input plaintext
2. Input cheating text
3. Does the cheating text contain all of the different characters in the plaintext?
   - Yes: Generate the character's position table (CPT) of the cheating text
   - No: Randomly generate an ID for the cheating text
4. Generate plaintext index file (PIF) by random
5. Compress the PIF
6. Encrypt the compressed PIF
7. Put the encrypted ID in the head of the encrypted and compressed PIF
8. Send out the compressed and encrypted PIF and the cheating text involving an ID

Receiver

(CDES, Lin & Lee, 1998 [1])
Decrypt the encrypted ID in the given PIF

Does find out the corresponding cheating text?

Yes

Decrpt the given PIF

Generate the character’s position table (CPT) of the cheating text

Decompress the given PIF

Use the CPT and the PIF to reverse the original plaintext

Plaintext output

No, wait the correct cheating text to come

(CDES, Lin & Lee, 1998 [1])

Receiver

2-nd key

1-st key

No
Input

Plaintext : Cat is my pet.
{C, a, t, i, s, m, y, p, e, , □}

Cheating text : Computer security is important.
{C, o, m, p, u, t, e, r, s, c, i, y, a, n, , □}

<table>
<thead>
<tr>
<th>Character</th>
<th>Position record</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1</td>
</tr>
<tr>
<td>o</td>
<td>2, 25</td>
</tr>
<tr>
<td>m</td>
<td>3, 23</td>
</tr>
<tr>
<td>p</td>
<td>4, 24</td>
</tr>
<tr>
<td>u</td>
<td>5, 13</td>
</tr>
<tr>
<td>t</td>
<td>6, 16, 27, 30</td>
</tr>
<tr>
<td>e</td>
<td>7, 11</td>
</tr>
<tr>
<td>r</td>
<td>8, 14, 26</td>
</tr>
<tr>
<td>s</td>
<td>10, 20</td>
</tr>
<tr>
<td>c</td>
<td>12</td>
</tr>
<tr>
<td>i</td>
<td>15, 19, 22</td>
</tr>
<tr>
<td>y</td>
<td>17</td>
</tr>
<tr>
<td>a</td>
<td>28</td>
</tr>
<tr>
<td>n</td>
<td>29</td>
</tr>
<tr>
<td>.</td>
<td>31</td>
</tr>
<tr>
<td>□</td>
<td>9, 18, 21</td>
</tr>
</tbody>
</table>

Table 1. Characters Position Table (CPT)
System model
Do you want to have a coffee with me?

It looks OK!

Eavesdropper

Cheating text

PIF

Sent out via Email

We should meet at 9 pm.

The proposal is based on Confused Document Encrypting Scheme
System model

Compression module
- Compress the PIF

Encryption module
- Encrypt the compressed PIF

Image-Hiding module
- Hiding the PIF in image

CDES module
- Plaintext
- Cheating text
- CPT generated
- PIF generated

Send out via E-mail
• Sender:
  – Generate the CPT by the cheating text, and the plaintext will generate PIF by CPT.

• Receiver:
  – Use the CPT and the PIF to reverse the original plaintext.
Compression Module

- It provides compression/decompression for the plaintext index file (PIF), because the PIF size will be large.

- It uses the LZMA algorithm.
Data Compression

- Reduce the data size.
- Decrease transmission time.
- Increase security of data.
LZMA algorithm

- Dictionary coding
  - Dictionary size increased $\rightarrow$ Higher compression rate $\uparrow$ and lower speed $\downarrow$

Example:
ABCDEFBCGXY

1. \{AB,BC,CD,DE,EF\} will be added to the dictionary and translated to a smallest unique-code.

2. Later, if BC has been stored in the dictionary, so it will be translated to a smallest unique-code, and add the strings \{BCG,XY\} to the dictionary.
You’ll remember what you had seen before.

Keep in mind (The dictionary)

What I’ve seen
Cat 1
Dog 2
Tree 3
• Encrypt the plaintext index file (PIF)

• It uses the Blowfish algorithm
In our case, we assume the keys already delivered securely.

**Cryptography**

- **M** (Plaintext) = original message
- **C** (Cipher text) = encrypted message
- **E_k** (M) = Encryption function
- **D_k** (C) = Decryption function

**Diagram:**

- **(Encryptor)**
  - **E**
  - **KEY1**
  - **M**
  - **C = E_k1(M)**

- **(Hacker)**
  - **Public area**
  - **C**

- **(Decryptor)**
  - **D**
  - **KEY2**
  - **M = D_k2(C)**

- **Private area**
Image-hiding Module

- It provides image hiding for the plaintext index file (PIF).

- Uses the "JPHS" (open source software).

- JPHS also uses the Blowfish algorithm to encrypt the PIF in an image.
Implementation
Environment

- Windows XP SP3
- Visual Studio 2005
- Mozilla Thunderbird 3 (3.0.4)
Do you want to have a coffee with me?

We should meet at 9 pm.
Receive a mail

Test on 7/3 16:51

from Weishin Pan <weishin.pan@gmail.com>
to magicpanx@gmail.com
date Sat, Jul 3, 2010 at 4:52 PM
subject Test on 7/3 16:51
mailed-by gmail.com
signed-by gmail.com

Do you want to have a coffee with me? → Cheating text

National Chiao Tung University
Graduate Institute of Communication Engineering
Next Generation Telephony Laboratory (R409)
Weishin Pan :)

Plaintext index file embedded
Flow chart (Sender)

1. Read the plaintext
2. Read the cheating text
3. Generate the CPT of the cheating text
4. Generate the plaintext index file
5. Compress the PIF
6. Encrypt the compressed PIF
7. Hiding the PIF in image
8. Compose a mail to receiver
9. From: weishin.pan@gmail.com
10. To: magicpanx@gmail.com
11. Subject: Hello, magicpanx!
12. Attached file: ncnu.jpg
   Cheating text
   PIF embedded

Do you want to have a coffee with me?
We should meet at 9 pm.

From: weishin.pan@gmail.com
To: magicpanx@gmail.com
Subject: Hello, magicpanx!

Do you want to have a coffee with me?
Attached file: ncnu.jpg

Cheating text

PIF embedded

Read the cheating text

Seek PIF from the attach image

Decrypt the PIF

Generate the CPT of the cheating text

Decompress the PIF

Using the CPT and the PIF to reverse the original plaintext

Plaintext output
Conclusion

• Increase the security in email services.

• Through the behavior observed in chatting, the image hiding technique is applied to CDES for hide the PIF in an emoticon or a picture, which looks meaningful.
Future work

• The framework can use for instant message (IM), like Windows Live Messenger, Yahoo Messenger in the future work.
  – Because people use the emoticon and photo in the chat, it has been a common behavior.
Thank you for your listening