College of Computer and Information Science
Northeastern University
CSG252: Cryptography and Communication Security

Fall 2004

Problem Set 3 (due October 26, 2004)
Submit electronically to ati@ccs.neu.edu and CC me.

Problem 1: Excel Worksheet Protection
http://chicago.sourceforge.net/devel/docs/excel/encrypt.html

MS Excel uses a very simple scheme to protect a sheet from modification. An Excel sheet can be protected using the menu: Tools -> Protection -> Protect Sheet.

Protection Scheme: the protection password is hashed using a simple scheme then stored in the file. Let the password be $p_1p_2\ldots p_k$. The hash H is computed as follows:

- $H = (p_1 <<_{15} 1) \oplus (p_2 <<_{15} 2) \oplus (p_3 <<_{15} 3) \oplus \ldots (p_k <<_{15} k) \oplus k \oplus 0xCE4B$.
- H is a 16 bits word (2 bytes) and is stored in the file.
- To unprotect the sheet Excel requests a password, computes its hash and verifies that it equals the stored hash.

$a <<_{15} b$ denotes shifting (w/ rotation) the ascii code of character 'a' to the left 'b' times using only the lowest 15 bits and rotating the highest bits (exceeding 15) to the right. Example: 'j' = 0x6A; 'a' $<<_{15}$ 10 = 0x2803.

Example: password "test" of 4 characters leads to hash 0xCBEB.

- a -> 0x61 << 1 == 0x00C2
- b -> 0x62 << 2 == 0x0188
- c -> 0x63 << 3 == 0x0318
- d -> 0x64 << 4 == 0x0640
- e -> 0x65 << 5 == 0x0CA0
- f -> 0x66 << 6 == 0x1980
- g -> 0x67 << 7 == 0x3380
- h -> 0x68 << 8 == 0x6800
- i -> 0x69 << 9 == 0x5201 (unrotated: 0xD200)
- j -> 0x6A << 10 == 0x2803 (unrotated: 0x1A800)

H = 0x00C2 ⊕ 0x0188 ⊕ 0x0318 ⊕ 0x0640 ⊕ 0x0CA0 ⊕ 0x1980 ⊕ 0x3380 ⊕ 0x6800 ⊕ 0x5201 ⊕ 0x2803 ⊕ 0x000A ⊕ 0xCE4B = 0xFEF1.

1. Protect an Excel using password 'test'. What is the value of the hash H?
2. Use a Hex Editor (such as: Hex Edit: http://www.expertcomsoft.com/) to open the file you protected.

   a. Identify where the 2 bytes of the hashed password are located? Hint: remember that the there are two system formats: Little-Endian and Big-Endian (http://www.webopedia.com/TERM/b/big_endian.html). Which one is used by Excel on PC?
   b. How can you identify the location of where the hashed password is stored in any file? Hint: look at the sequence of characters preceding/following the hash. What can you say about it? How about the password of other worksheets.
3. Find two passwords different from 'test' that can allow you to unprotect a sheet protected with password 'test'. You might find helpful to use the table of ascii encoding: http://www.cplusplus.com/doc/papers/ascii.html.
4. Discuss why the excel protection scheme is weak.
5. **Bonus points:** Write a program that automatically unprotect all sheets in an excel file.

Problem 2: Textbook problem 3.14.

Problem 3: Excel Workbook Encryption.

1. Write the pseudo-code for Excel Workbook Encryption assuming the following documentation is correct: **http://chicago.sourceforge.net/devel/docs/excel/encrypt.html).**
2. Discuss why MS Excel Workbook Encryption is weak.
3. Describe how you can unprotect an Excel Workbook.
4. How much time your algorithm would take if assuming *m* units of time for running MD5 hash, and *r* units for encrypting 128 bits using RC4 algorithm.
5. Discuss how this protection could be made stronger and what could be the reasons for its weak original design.

For MS students only:

Problem 4: Propose an algorithm for inversing a number mod m.

Problem 5: Textbook 3.10.

For PhD Students only:

Problem 6: Textbook 3.8.

Problem 7: Textbook 3.10.

Problem 8: Textbook 3.13.