C&O 487/687: Assignment #2

Due date: October 18, 2022 8:59:59pm (Thursday)

Note: Some questions use randomization to customize to you specifically. Please include your max-8-character UW user id (`j2tham`) at the beginning of your answer so we can look up your custom solution.

1. [10 marks] **Linear cryptanalysis**

   Please include your max-8-character UW user id (`j2tham`) at the beginning of your answer so we can look up your custom solution.

   This problem uses the simplified cipher described in Section 2 of "A Tutorial on Linear and Differential Cryptanalysis" by Howard M. Heys, available at `http://www.engr.mun.ca/~howard/Research/Papers/ldc_tutorial.html`. We refer to this cipher as the "Heys cipher".

   For the purposes of this problem, each student has a fixed, unknown 80-bit key. You will be carrying out a known-plaintext attack against the Heys cipher using linear cryptanalysis, using a set of 20,000 distinct random plaintext-ciphertext pairs. You can download your plaintexts and ciphertexts (unique to you) at the following addresses:

   `https://www.math.uwaterloo.ca/~dstebila/as/as.cgi/download/co487_f22/a2q1plaintexts.txt`

   `https://www.math.uwaterloo.ca/~dstebila/as/as.cgi/download/co487_f22/a2q1ciphertexts.txt`

   The format of the files is that the nth line of the ciphertext file equals the encryption of the nth line of the plaintext file under your secret key.

   For the programming aspects of questions 1.a.i, 1.b, and 1.d.ii, you may work with a partner to do the programming and you may submit the same computer program source code, but you must submit your own write-up and explanations for the non-programming parts of those questions. Please indicate in your submission who you worked with.

   (a) [2 marks] Using the linear approximation $U_{4,6} \oplus U_{4,8} \oplus U_{4,14} \oplus U_{4,16} \oplus P_5 \oplus P_7 \oplus P_8 \cong 0$, Carol guesses that the target partial subkey $[K_{5,5}, K_{5,6}, K_{5,7}, K_{5,8}, K_{5,13}, K_{5,14}, K_{5,15}, K_{5,16}]$ has the value $[0, 1, 1, 1, 0, 1, 1, 0]$. Note that Carol's guess is not necessarily correct! Do one of the following:

      i. Determine the magnitude of the bias for this partial subkey value over your twenty thousand plaintext/ciphertext pairs, using a computer program, and provide the source code for your program, or:

      ii. Determine the magnitude of the bias for this partial subkey value over your first *ten* plaintext/ciphertext pairs, without using a computer program, and show your work.

   (b) [3 marks] Find the value of the partial subkey $[K_{5,5}, K_{5,6}, K_{5,7}, K_{5,8}, K_{5,13}, K_{5,14}, K_{5,15}, K_{5,16}]$ for your key, by calculating the target partial subkey which yields the largest magnitude of bias over your 20,000 plaintext/ciphertext pairs. You will almost certainly need a computer program for this task; provide the source listing for any computer code that you or your collaborators write.

   (c) [2 marks] By using Table 4 in the tutorial, compute the bias in each of the following individual S-box linear approximations:

   $$S_{11} : X_1 \oplus X_4 \cong Y_1$$

   $$S_{13} : X_1 \oplus X_4 \cong Y_1$$

   $$S_{21} : X_1 \oplus X_3 \cong Y_2$$

   $$S_{32} : X_1 \cong Y_1 \oplus Y_2 \oplus Y_3 \oplus Y_4$$

   Then, combine these to find a linear approximation of the first three rounds of the Heys cipher, and calculate the theoretical magnitude of its bias.

(d) [3 marks] Do *one* of the following:

i. Using the linear approximation from part (c), determine the entire subkey K_5. You will almost certainly need a computer program for this task; provide the source listing for any computer code that you or your collaborators write, or:

ii. Using the (incorrect) guess $K_5 = [1100011111100110]$ for the fifth subkey, determine the magnitude of the bias of the approximation from part (c) for this subkey over your first ten plaintext/ciphertext pairs, without using a computer program, and show your work.

(e) [2 bonus marks] A small amount of extra credit is available if you can determine any additional key bits.

2. [6 marks] **Differential cryptanalysis**

Please include your max-8-character UW user id (`j2tham`) at the beginning of your answer so we can look up your custom solution.

This problem uses the simplified cipher described in Section 2 of "A Tutorial on Linear and Differential Cryptanalysis" by Howard M. Heys, available at `http://www.engr.mun.ca/~howard/Research/Papers/ldc_tutorial.html`. We refer to this cipher as the "Heys cipher".

For the purposes of this problem, each student has a fixed, unknown 80-bit key. You will be carrying out a chosen-plaintext attack against the Heys cipher using differential cryptanalysis, using a set of 5,000 distinct random plaintext-ciphertext pairs. You can download your plaintexts and ciphertexts (unique to you) at the following address:

`https://www.math.uwaterloo.ca/~dstebila/as/as.cgi/download/co487_f22/a2q2ciphertexts.txt`

Each line in the file has the following format:

$$aaaaaaaaaaaaaaaa, bbbbbbbbbbbbbbbb, cccccccccccccccc, dddddddddddddddd$$

where:

- *aaaaaaaaaaaaaaaa* is the plaintext X', in bits.
- *bbbbbbbbbbbbbbbb* is the plaintext $X'' = X' \oplus \Delta X$.
- *cccccccccccccccc* is the ciphertext Y' (encryption of X'), encrypted under your key.
- *dddddddddddddddd* is the ciphertext Y'' (encryption of X''), encrypted under your key.

Each plaintext pair uses the same fixed difference $\Delta X = 0000\ 1011\ 0000\ 0000$.

Note that your keys/plaintexts/ciphertexts for this problem are **NOT** the same as in Problem 1.

(a) [3 marks] Do *one* of the following:

i. Find the value of the partial subkey $[K_{5,5}, K_{5,6}, K_{5,7}, K_{5,8}, K_{5,13}, K_{5,14}, K_{5,15}, K_{5,16}]$ for your key, by calculating the target partial subkey which yields the largest number of plaintext/ciphertext pairs satisfying the equation $\Delta U_4 = [0000\ 0110\ 0000\ 0110]$. You will almost certainly need a computer program for this task; provide the source listing for any computer code that you or your collaborators write.

ii. Using the (incorrect) guess $K_5 = [1100\ 0111\ 1110\ 0110]$ for the fifth subkey, compute the value of ΔU_4 for each of the first ten plaintext/ciphertext pairs from the list for your key, without using a computer program, and show your work.

(b) [3 marks] Using the following difference pairs:

$$S_{11}: (\Delta X, \Delta Y) \mapsto (A, 8)$$
$$S_{13}: (\Delta X, \Delta Y) \mapsto (A, 8)$$
$$S_{21}: (\Delta X, \Delta Y) \mapsto (A, 8)$$
$$S_{31}: (\Delta X, \Delta Y) \mapsto (8, B)$$

find the corresponding differential characteristic and calculate its probability. Explain how you would use this differential to find K_5.

3. [8 marks] **Block cipher modes of operation**

   Consider the *stateful block chaining* (SBC) block cipher mode of operation. In addition to a uniformly random 128-bit key k, Alice keeps track of an initialization vector n, which is initialized to a uniformly random 128-bit string. With this mode of operation, we will allow Alice to encrypt multiple multi-block messages without needing to send a new IV for each separate message; she sets the IV for the next message based on some calculation from the previous message (hence "stateful").

   Alice encrypts a message m as follows:

   (a) Pad m so that its length is a multiple of 128 (the details of the padding scheme are not important for this question).

   (b) Divide m into 128-bit blocks $m_0, m_1, \ldots, m_{\ell-1}$, so that $m = m_0 \| m_1 \| \ldots \| m_{\ell-1}$.

   (c) Set $c_0 \leftarrow \text{AES}_k(m_0 \oplus \mathsf{IV})$.

   (d) For $i = 1, \ldots, \ell - 1$, set $c_i \leftarrow \text{AES}_k(m_i \oplus c_{i-1})$.

   (e) Save $\mathsf{IV} \leftarrow c_{\ell-1}$.

   (f) Output ciphertext $c_0 \| c_1 \| \ldots \| c_{\ell-1}$.

   Alice includes the initialization value n in her first message to Bob; in subsequent messages, this is not necessary since a previous ciphertext block is used as the IV.

   To attack SBC mode, we will give the adversary additional powers, in what we will call a *chosen plaintext attack + prefixes* attack. In a CPA+P attack, the adversary has the same capabilities as in a chosen plaintext attack, but they additionally can obtain encryptions of the challenge message with arbitrary prefixes. For example, if the challenge message is m^*, then the adversary can query the encryption of $0 \| m^*$, $111100111 \| m^*$, or even just m^*.

   In this question, you will devise a chosen plaintext + prefixes attack on AES in SBC mode.

   (a) [2 marks] Suppose that a message $m = m_0 \| \ldots \| m_{\ell-1}$ is encrypted to the ciphertext $c = c_0 \| \ldots \| c_{\ell-1}$ using SBC mode with IV n. Suppose that B is the next block encrypted after m, and let C be the encryption of B. For what value of B will $C = c_0$?

   (b) [2 marks] With the same setup as in part (a), suppose that c is the challenge ciphertext and n is known to the adversary. Explain how the adversary can check if $m'_0 = m_0$ for any 128-bit block m'_0.

   (c) [4 marks] Describe a chosen plaintext attack + prefixes attack which totally breaks AES in SBC mode.

   Hint: Query the encryption of $0^{127} \| m$.

4. [6 marks] **Exploring preimage resistance**

   We say a hash function H is *preimage resistant* for messages of length m if, given $y = H(x)$ for $x \in_R \{0,1\}^m$ (x chosen uniformly at random from $\{0,1\}^m$), it is computationally infeasible to find (with non-negligible probability of success) any input z such that $H(z) = y$.

   We say a hash function H is *second preimage resistant* for messages of length m if, given an input $x \in_R \{0,1\}^m$, it is computationally infeasible (with non-negligible probability of success) to find a second input $x' \neq x$ such that $H(x) = H(x')$.

   (a) [3 marks] Suppose that a second preimage resistant hash function $H : \{0,1\}^* \to \{0,1\}^n$ is "somewhat uniform", meaning that all hash values have approximately the same number of preimages. Show that H is preimage resistant for messages of length m, when $m \geq 2n$.

(b) [3 marks] Give an example of a second preimage resistant hash function where, given a hash value $y \in_R \{0,1\}^n$, it is computationally feasible to find an input x such that $H(x) = y$, with non-negligible probability of success.

Hint: Start with a second preimage resistant hash function $G\colon \{0,1\}^* \to \{0,1\}^{n-1}$. From G construct another hash function which has the desired property.

5. [5 marks] **A weird hash function**

For bits $x, y \in \{0,1\}$, the operator $\odot$ is defined as follows:

$$x \odot y = \begin{cases} 1 & \text{if } x \neq y \\ 0 & \text{otherwise} \end{cases}.$$

For bit strings $x, y \in \{0,1\}^n$, we apply the operation bitwise. For example, if $x = 110$ and $y = 010$, then $x \odot y = 100$.

Let $m = 128$. Suppose that $f : \{0,1\}^m \to \{0,1\}^m$ is a preimage resistant bijection. Define $h : \{0,1\}^{2m} \to \{0,1\}^m$ as follows. Given $x \in \{0,1\}^{2m}$, write

$$x = x' \,||\, x''$$

where $x', x'' \in \{0,1\}^m$. Then define

$$h(x) = f(x' \odot x'').$$

(a) [3 marks] Prove that h is not second preimage resistant.

(b) [2 marks] Prove that h is preimage resistant for messages of length $2m$.

Academic integrity rules

You should make an effort to solve all the problems on your own. You are also welcome to collaborate on assignments with other students in this course. However, solutions must be written up by yourself. If you do collaborate, please acknowledge your collaborators in the write-up for each problem. *If you obtain a solution with help from a book, paper, a website, or any other source, please acknowledge your source. You are not permitted to solicit help from other online bulletin boards, chat groups, newsgroups, or solutions from previous offerings of the course.*

Due date

The assignment is due via Crowdmark by 8:59:59pm on October 18, 2022 8:59:59pm. Late assignments will not be accepted.

Office hours

Office hours will take place online via the Gather.town platform; see the link on LEARN under Contents → Course Information → Office hours.

- Monday September 26 11am–12pm
- Thursday September 29 1–2pm
- Monday October 3 11am–12pm

- Tuesday October 4 11am–12pm
- Wednesday October 5 11am–12pm
- Wednesday October 5 2–3pm
- Wednesday October 5 4–5pm
- Thursday October 6 1–2pm

Changelog

- Fri. Sep. 23: assignment posted
- Tue. Sep. 27: due date extended to October 18 at 8:59:59pm
- Tue. Oct. 4: question 4(b) revised: removed condition on "somewhat uniform"