CS 6324: Information Security **Spring 2021**

Homework #1

Due date & time: 11:59pm CST on February 12, 2021. Submit to eLearning by the due time.

Additional Instructions: The submitted homework must be typed. Using LaTeX is recommended, but not required.

Problem 1 (10 pts) Confidentiality, Integrity, Availability.

- (3 pts) State what is Confidentiality, Integrity, and Availability.
- (3 pts) For each, give two examples where they are violated.
- (4 pts) Identify two computer security control measures on your computer(s). Which of the three properties Confidentiality, Integrity, and Availability do they aim at providing? What kinds of adversaries they **cannot** defend against?

Problem 2 (30 pts) Probability Review.

1. (10 pts) We roll two fair 6-sided dice.
   (a) Find the probability that doubles are rolled.
   (b) Given that the roll results in a sum of 6 or less, find the conditional probability that doubles are rolled.
   (c) Find the probability that that larger of the two die's outcomes is at least 4.
   (d) Given that the two dice land on different numbers, find the conditional probability that at least one die roll is a 1.
   (e) Let X be the event that the first die results in an odd number, and Y be the event that the rolled sum is an even number. Compute $\mathbf{Pr}[X], \mathbf{Pr}[Y], \mathbf{Pr}[X \wedge Y]$. Are X and Y independent?
   (f) Let X be the event that the first die results in a multiple of 2, and Y be the event that the rolled sum is a multiple of 2. Compute $\mathbf{Pr}[X], \mathbf{Pr}[Y], \mathbf{Pr}[X \wedge Y]$. Are X are Y independent?
   (g) Let X be the event that the first die results in a multiple of 3, and Y be the event that the rolled sum is a multiple of 2. Compute $\mathbf{Pr}[X|Y]$ and $\mathbf{Pr}[Y|X]$. Are X are Y independent?

2. (5 pts) Let X denote the event that a student pass the midterm exam in a course. And Y denote the event that the student pass the final exam. We know that $\mathbf{Pr}[X] = 0.75$, $\mathbf{Pr}[Y] = 0.8$, and $\mathbf{Pr}[Y|X] = 0.9$. (The probability that a student will pass the final exam given that he or she has passed the midterm exam is 0.9.) Compute $\mathbf{Pr}[X \wedge Y]$, $\mathbf{Pr}[\neg X|Y]$, and $\mathbf{Pr}[Y|\neg X]$. Are X and Y independent?

3. (6 pts) There are 78 qualified applicants for teaching positions in an elementary school, of which some have at least five years' teaching experience and some have not, some are married and some are single, with the exact breakdown being

	Married	Single
At least five years teaching experience	18	12
Less than five years teaching experience	30	18

   (a) The order in which the applicants are interviewed is random. M is the event that the first applicant interviewed is married and F is the event that the first applicant interviewed has at least five years teaching experience. Find the following probabilities: $\mathbf{Pr}[M], \mathbf{Pr}[F], \mathbf{Pr}[M \wedge F], \mathbf{Pr}[M|F], \mathbf{Pr}[F|M]$. Are M and F independent?

   (b) Suppose that there is only one opening in the third grade, and each applicant with at least five year experience has *twice* the chance of an applicant with less than five year experience. Let U denote the event that the job goes to one of the single applicants, and V the event that it goes to an applicant with less than five year experience. Find the following probabilities: $\mathbf{Pr}[U], \mathbf{Pr}[V], \mathbf{Pr}[U \wedge V], \mathbf{Pr}[U|V], \mathbf{Pr}[V|U]$. Are U and V independent?

4. (9 pts) A test for a certain rare disease is assumed to be correct 95% of the time. Let S denote the event that a person has the disease, and T denote the event that the test result is positive. We have $\mathbf{Pr}[T|S] = 0.95$ and $\mathbf{Pr}[\neg T|\neg S] = 0.95$. Assume that $\mathbf{Pr}[S] = 0.001$, i.e., the probability that a randomly drawn person has the disease is 0.001.

   (a) Compute $\mathbf{Pr}[S|T]$, the probability that one has the disease if one is tested positive.

   (b) Assume that one improves the test so that $\mathbf{Pr}[T|S] = 0.998$, i.e., if one has the disease, the test will come out positive with probability 0.998. Other things remain unchanged. Compute $\mathbf{Pr}[S|T]$.

   (c) Assume that another improvement on the test improves $\mathbf{Pr}[\neg T|\neg S] = 0.998$, but have $\mathbf{Pr}[T|S]$ remain at 0.95. Compute $\mathbf{Pr}[S|T]$.

Problem 3 (8 pts) Cryptanalysis Concepts.

- (3 pts) Explain what do ciphertext-only attacks, known-plaintext attack, and chosen plaintext attack mean?
- (2 pts) What attack can be used to break the substitution cipher under a ciphertext-only attack? Explain the simplest way to break it under a known-plaintext attack?
- (3 pts) Explain how one may be able to carry out a known-plaintext attack against the wireless encryption between the laptop used by the target user and a wireless access point. Explain how one may be able to carry out a chosen-plaintext attack.

Problem 4 (8 pts) Consider the following "Double Vigenère encryption". We choose two random keys K_1 and K_2 of lengths ℓ and $\ell + 1$. To encrypt a message, we first use key K_1 to encrypt in the Vigenère fashion, then use K_2 to encrypt.

- (4 pts) Describe how to break this encryption scheme under a ciphertext only attack. Does this double encryption offer increased level of security over Vigenère encryption against a ciphertext only attack?
- (4 pts) Suppose that we know that $10 \le \ell \le 19$. Given a ciphertext of length 50 and its corresponding plaintext, describe how to recover the keys so that any other message encrypted under the same key pair can be decrypted.

Problem 5 **(14 pts)** Consider the following enhancement of the Vigenère cipher. We assume that the plaintext is a case-insensitive English text using only the 26 letters (without space or any other symbol). To encrypt a plaintext of length n, one first uniformly randomly generates a string over the alphabet $[A..Z]$ of length 17, and then inserts this string into the beginning of the plaintext. That is, we first construct a string $x = x_1 x_2 \ldots x_{n+17}$, such that $x_1 \cdots x_{17}$ is the string we have generated, and $x_{18} \cdots x_{n+17}$ is the original plaintext string. We then construct a string y as follows: $y_i = x_i$ for $1 \le i \le 17$, and for $i \in [18, n+17]$, y_i is the result of using y_{i-17} to encrypt x_i; that is, when the x_i's and y_i's are treated as numbers in $[0..25]$, we have $y_i \leftarrow ((x_i + y_{i-17}) \mod 26)$. We then apply the Vigenère cipher to the string y, while making sure that the key length is not a multiple of 17.

- Implement the encryption algorithm and decryption algorithm for this cipher in a programming language of your choice. Include the core part of your code.
- Choose a key, a plaintext, and run the encryption code multiple times, and ensure that the decryption results in the original plaintext. Include the key, the plaintext, and 3 ciphertexts.
- Write the Pseudo-code to recover enough information from a known (plaintext,ciphertext) pair to decrypt other messages encrypted under the same key. That is, the pseudo-code takes three inputs (M_1, C_1, C_2), where C_1 is a ciphertext of M_1, and outputs M_2, the message encrypted in C_2.
- How to effectively attack this cipher in a ciphertext only attack?

Hint: You may want to do this question last.

Problem 6 **(10 pts)** Consider an example of encrypting the result of a 6-side dice (i.e., $M \in [1..6]$), as follows. Uniformly randomly chooses $K \in [1..6]$, ciphertext is $C = (M * K) \mod 13$. The ciphertext space is thus $[1..12]$. We have $\mathbf{Pr}[\mathsf{PT} = 1] = \mathbf{Pr}[\mathsf{PT} = 2] = \mathbf{Pr}[\mathsf{PT} = 3] = \cdots = \mathbf{Pr}[\mathsf{PT} = 6] = 1/6$, and we use a vector notation $\mathbf{Pr}[\mathsf{PT}] = \langle 1/6, 1/6, \ldots, 1/6 \rangle$ to denote this.

- Assume that you stole a glance at the dice value and saw that there are many dots on it, and hence are quite certain that M is either a 5 or a 6. You then learned the ciphertext of the encrypted dice value. Under which ciphtertext value(s) can you learn the value M.

  Hint: You may want to start by writing out the 6 by 6 table of the ciphertext for each possible combination of plaintext and key.
- Compute $\mathbf{Pr}[\mathsf{PT}|\mathsf{CT} = i]$, for each $i \in [1..12]$, similar to the one for $\mathbf{Pr}[\mathsf{PT}]$ given above.
- Show that if K is uniformly randomly chosen from $[1..12]$, then this cipher provides perfect secrecy.

Problem 7 (5 pts) Consider the following way of using the Vigenere cipher to send one encrypted message. The possible plaintexts are English texts of length 100. The key is a random string of length 50. Show that this does not satisfy perfect secrecy by finding two plaintext messages M_1, M_2 and a ciphertext message C such that:

$$\mathbf{Pr}[\mathsf{CT} = C \mid \mathsf{PT} = M_1] \neq \mathbf{Pr}[\mathsf{CT} = C \mid \mathsf{PT} = M_2]$$

Problem 8 (5 pts) Prove that for any cipher that offers perfect secrecy, the number of the possible keys must be at least as large as the number of plaintexts.

Problem 9 (10 pts) Implement the following encryption/decryption function, which uses the RC4 stream cipher.

- byte[] encrypt(byte[] pt, byte[] key)
- byte[] decrypt(byte[] ct, byte[] key)

You should implement RC4 algorithm yourself. Google to find out the algorithm. You can assume that the key is an array of length between 16 and 32 bytes. You need to use a 256-bit (32-byte) initial vector so that when one invokes the encrypt function with the same pair of plaintext and key twice, with high probability the resulting ciphertexts are different.

You can choose any programming language to do this problem. You need to figure out what library function to call to generate a random IV. Note that the random IV is required to be unpredictable. This call to generate IV should be the only library call. You should drop the first 3072 bytes of RC4's output as recommended. You should run your code to verify that decrypt(encrypt(pt, key), key) = pt.

Include your code in the HW submission, and provide information regarding the library function you use to generate the random IV.