STUDENT VERSION

May 15, 2017

Encryption: Security Through Mathematics

Center for Science of Information, A National Science Foundation Science & Technology Center `http://soihub.org`

Questions on *Encryption: Security Through Mathematics*

1. Which of the following is/are true:
   - ◯ A continuous memoryless channel takes continuous inputs and continuous outputs, and is indexed by a continuous time index.
   - ◯ C(P) is non-decreasing, concave and left-continuous for any real function $\kappa(x)$.
   - ◯ $\frac{1}{n}\sum_{i=1}^{n}\kappa(x_i) \leq P$ in the definition of (κ, P) implies that the input distribution here is restricted to discrete random variable X.
2. Which of the following about the converse proof is/are true:
   - ◯ $h(\mathbf{Y}|\mathbf{X}) = \sum_{i=1}^{n} h(Y_i|X_i)$ due to the retention property of the channel.
   - ◯ We have to show $I(X_i;Y_i) \leq C(P)$ before we show $\sum_{i=1}^{n} I(X_i;Y_i) \leq nC(P)$.
   - ◯ The concavity of mutual information $\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \leq I(X;Y)$ is based on $X \sim \frac{1}{n}\sum_{i=1}^{n} F_i(x)$.
   - ◯ The concavity of mutual information $\frac{1}{n}\sum_{i=1}^{n} I(X_i;Y_i) \leq I(X;Y)$ is based on $X \sim \frac{1}{n}\sum_{i=1}^{n} X_i$.
3. Is this true (Y/N): It is difficult to formulate the notion of jointly typicality as in discrete case because the input distribution F(x) may not have a pdf.
   - ◯ Y
   - ◯ N
4. Is this true (Y/N): The random coding scheme requires the sequence length n to be sufficiently large.
   - ◯ Y
   - ◯ N
5. Is this true (Y/N): In the random coding scheme, there is zero probability for a codeword to violate the constraint.
   - ◯ Y
   - ◯ N
6. Which of the following is/are true:

- ◯ The signal-to-noise ratio P/N of a memoryless Gaussian channel has to be greater than 1 so that its capacity is positive.
- ◯ If there is an input power constraint, we can transmit information through a memoryless Gaussian channel reliably at any finite rate.
- ◯ $h(Y|X) = h(Z|X)$ for Y = X+Z because conditioning on X, Y and Z determine each other. Therefore, given X, Y and Z carries the same amount of information.
- ◯ $h(Z|X) = h(Z)$ because Z is independent of X.

7. The capacity of the parallel Gaussian channels system is attained when:
   - ◯ The inputs are dependent of each other.
   - ◯ The inputs are Gaussian.
   - ◯ The output powers are allocated according to water-filling.
   - ◯ Every channel is used.
8. Which of the following is/are true:
   - ◯ The equivalent system is a system of Gaussian channels with decorrelated noise vector.
   - ◯ The capacity of a system of correlated Gaussian channels is attained when the inputs are mutually independent.
   - ◯ Multiplying an input vector $\mathbf{X}$ by Q and then transmitting it through the correlated Gaussian channels produce the same output as transmitting it through the equivalent channel.
9. Which of the following is true:
   - ◯ The orthonormal basis of signals bandlimited to $0, W$ is given by $\psi_i(t) = \sqrt{2W}\text{sinc}(2Wt - i), i \in (-W, W)$.
   - ◯ The orthonormal basis of signals bandlimited to $0, W$ is given by $\psi_i(t) = \frac{1}{\sqrt{2W}}\text{sinc}(2Wt - i), i \in (-\infty, \infty)$.
   - ◯ $\psi_i(t) i \in (-\infty, \infty)$ form an orthonormal because each $\psi_i(t)$ has energy 1 and $\psi_i(t), \psi_j(t)$ has cross-correlation function 0 at $\tau = 0$ for distinct i, j.
10. Which of the following is/are true:
    - ◯ The bandlimited white Gaussian channel is equivalent to a system of correlated Gaussian channels.
    - ◯ The noise power of the corresponding memoryless Gaussian channel is N_0 W.
    - ◯ $Z(\frac{i}{2W}) i \in (-\infty, \infty)$, i.e. the unfiltered $Z(t)$ with $t = \frac{i}{2W}$, are i.i.d. Gaussian random variables with zero mean.
    - ◯ If $X(t)$ has power constraint P, then X_i is constrained to $\frac{P}{2W}$.
11. Capacity of the bandlimited colored Gaussian channel:
    - ◯ $C(P) = \int_{-W}^{W} \log(1 + \frac{(-S_Z(f))^+}{S_Z(f)}) df$, where $\frac{1}{2}\int_{-W}^{W} (-S_Z(f))^+ df = P$.
    - ◯ $C(P) = \frac{1}{2}\int_{-W}^{W} 1 + \frac{(-S_Z(f))^+}{S_Z(f)} df$, where $\int_{-W}^{W} \log(-S_Z(f))^+ df = P$.
    - ◯ $C(P) = \frac{1}{2}\int_{-W}^{W} \log(1 + \frac{(-S_Z(f))^+}{S_Z(f)}) df$, where $\int_{-W}^{W} (-S_Z(f))^+ df = P$.
12. Which of the following is true:
    - ◯ The zero-mean Gaussian noise has the largest differential entropy among all the additive noises with the same correlation matrix.
    - ◯ The diagonal elements of the correlation matrix after diagonalization specify the power of the individual noise variables.
    - ◯ Data processing increases the information divergence between two random variables.

Vigenere Cipher

1. The Caesar cipher is a simplification of the Vigenere cipher, in which all characters are shifted by the same amount, the key letter. Use the Caesar cipher to encode the following message with the key letter "K":

   `Birds of a feather flock together.`

2. Use the Vigenere cipher to encode the following message using the keyword "key":

   `The quick brown fox jumps over the lazy dog.`

3. Suppose we wanted to crack the Vigenere cipher and we were given a long, encoded message M. Suppose a and b are two substrings in M which match. What are possibilities for the length of the key? What can we do once we know the length of the key?

Playfair Cipher

One way to develop a Playfair square is to use a five-by-five square and omit the letter J. That is, during encoding and decoding, $I = J$. Moreover, the first row(s) will be the keywords with redundant letters omitted, and then the remaining squares are filled in with the alphabet, in order. For example, the Playfair square for keyword "CRYPTOGRAPHY" is:

C	R	Y	P	T
O	G	A	H	B
D	E	F	I	K
L	M	N	Q	S
U	V	W	X	Z

4. What is the Playfair square for the keyword "BOILERMAKER" look like?

5. What is the encryption, with the previous Playfair square for key "BOILERMAKER", for the following phrase:

   `No free lunch`

Strengthening the Vigenere Cipher

Use the one-time pad "A penny saved is a penny earned" to encrypt the following messages:

6. `Practice makes perfect`

7. `Just use frequency analysis!`

Cryptography and Entropy Part 1

8. Is the Caesar Cipher perfectly secure?

9. Intuitively, is the Vigenere Cipher perfectly secure?

10. Even though using a one-time pad is perfectly secure, why might using the same one-time pad repeatedly be insecure?

Cryptography and Entropy Part 2

11. A scheme that is perfectly secure has
$$H(M|C) = H(M).$$
What relationship holds for all encryption schemes?

12. If an encryption scheme is perfectly secure, how do the number of keys and the number of possible messages relate? What does this say about perfect security in practice?

13. In the previous section, we determined that intuivitely, the Vigenere Cipher is not perfectly secure. Now prove mathematically, that both the Vigenere Cipher and the Playfair Cipher are not perfectly secure.

Cryptography and Entropy Part 3

14. Compute the unicity point of the Vigenere Cipher for the English language using keywords of length 3, and only using frequency analysis. ($H = 4.2$)

15. How do the unicity points relate for the Caesar Cipher and the Vigenere Cipher? Why is this the case?

One-Way Functions

16. Define a one-way function. Why might it be useful for encryption?

17. Determine a method which uses multiplication as a one-way function.

Discrete Logarithms, Diffie-Hellman Key Exchange

18. Compute the product of the two primes, $3253 \cdot 3559$. Factor the number 1927. Which problem was more difficult?

19. Compute $5^{32} \pmod{33}$. Determine p for which

$$14^p \equiv 5 \pmod{17}.$$

Which problem was more difficult?

20. What is the purpose of the Diffie-Hellman Key Exchange? That is, what does the Diffie-Hellman Key Exchange accomplish?

21. Derive a key-exchange system for three parties, in a similar way to the Diffie-Hellman exchange.

The RSA Algorithm Under the Hood Part 1: Prime Numbers, Euclidean Algorithm

22. In the corresponding lecture, it was proven that there are infinitely many prime numbers. In the same idea of the proof, given a positive integer n, determine n consecutive composite numbers.

The RSA Algorithm Under the Hood Part 2: Prime Numbers, Euclidean Algorithm

23. Use the Euclidean Algorithm to find the greatest common divisor of 3924 and 56232

24. Use the Euclidean Algorithm to determine whether 724 and 813 are relatively prime.

The RSA Algorithm Under the Hood Part 3: Greatest Common Divisor and Fermat's Little Theorem

25. Use the Euclidean Algorithm to determine a and b such that

$$2232a + 1645b = 1.$$

26. Use the Euclidean Algorithm to determine the greatest common divisor, g, of 31824 and 864, and a and b such that

$$31824a + 864b = g.$$

The RSA Algorithm Under the Hood Part 4

27. Fermat's Little Theorem states that for p prime and a relatively prime to p,

$$a^p \equiv a \pmod{p}.$$

Find a counterexample if p can be any integer and a is not relatively prime to p.

28. The RSA encryption algorithm requires finding b, given an integer a, such that

$$ab \equiv 1 \pmod{m},$$

for some modulo m. How can this be achieved?

29. Given $a = 17$ and $m = 329$, find b such that

$$ab \equiv 1 \pmod{m}.$$

30. What type of encryption does RSA achieve? That is, what scenario is the RSA encryption algorithm useful?

The RSA Algorithm Under the Hood Part 5

In the RSA encryption algorithm, p, q are primes such that $n = pq$ and a, b are integers such that

$$ab \equiv 1 \pmod{(p-1)(q-1)}.$$

31. Suppose Alice has knowledge of all the variables and wants to implement the algorithm. Which variables does she publish?

32. Suppose Bob wants to encrypt message M and send it to Alice. What does he send?

33. How does Alice decrypt the received message M'?

34. Suppose an adversary intercepts message M'. What makes M' difficult to decrypt?

The RSA Algorithm Under the Hood Part 6

35. Compute by hand $5^{82} \pmod{3123}$.