Principles of Information Security

Assignment-1

1. (a) Consider You are designing a encryption scheme. Let say, the Message space is $\mathcal{M}$. What are the standard algorithms/components that you must have to define for your scheme. Explain these components. Explain their possible nature (Like whether they are probabilistic or deterministic) in case of perfectly secret private key encryption scheme.

   (b) Give formal specifications of the above mentioned components for at least two historical cipher. Make proper assumptions if needed.

2. (a) Explain Shift cipher and the Vigenere cipher in brief. Show how to use the Vigenere cipher for encryption of a word of length l. Is it possible to achieve perfect secrecy using Vigenere cipher in the above encryption? (you can make proper assumption about the key.) Prove your answer. Briefly Explain how someone can break Shift and Vigenere ciphers? Is there any case when Vigenere cipher is perfectly secret? Explain your answer.

   (b) Consider the Vigenere cipher over the lowercase English alphabet, where the key length can be anything from 8 to 12 characters. What is the size of the key space for this scheme?

   (c) Can you perform some modification in the standard version of Vigenere Cipher using the approach explained below? Is the Modified version completely secure, or can you break this Modified version? Justify your answer.
   Approach: For the Modified version of Vigenere cipher, one approach would be to Consider using multiple mono-alphabetic substitution ciphers, instead of using multiple shift ciphers. That is, the key consists of t random permutations of the alphabet, and the plaintext characters in positions i; t + i; 2t + i and so on are encrypted using the ith permutation.

3. Show that Shift, Substitution, Vigenere Ciphers are all trivial to break using a known-plaintext attack. (Assume only normal English words are being encrypted in each case.) how much known plaintext is needed to completely recover the key for each of the ciphers (without resorting to any statistics)?

4. Perfect Secrecy

   (a) Prove or refute: Every encryption scheme for which the size of the key space is equal to the size of the message space, and for which key are chosen uniformly from the key space, is perfectly secret.

   (b) Prove or refute: Consider a scheme for shift cipher where only a single character is encrypted.This scheme is perfectly secret.

(c) Prove or refute: For every encryption scheme that is perfectly secret it holds that for every distribution over the message space $\mathcal{M}$ every $m, m' \in \mathcal{M}$ and every $c \in \mathcal{C}$

$$\Pr[M = m | C = c] = \Pr[M = m' | C = c]$$

(d) Prove or refute: One time pad is a perfectly-secret encryption scheme.

(e) What is the largest plaintext space $\mathcal{M}$ you can find for which the mono-alphabetic substitution cipher provides perfect secrecy? (Note: $\mathcal{M}$ need not contain only valid English words.)

5. (a) Consider the Vigenere cipher over the lowercase English alphabet, where the key can have length 1 or length 2, each with 50 percent probability. Say the distribution over plain texts is Pr[M='aa'] = 0.4 and Pr[M='ab'] = 0.6. What is Pr[M='aa' | C='bb']? Express your answer to 4 decimal places.

   (b) Suppose a message space is of 5-bit strings. Consider the one-time pad over on this message space. Given Pr[M=00100] = 0.1 and Pr[M=11011] = 0.9. What is Pr[C=00000]? Calculate answer to 5 decimal places with a leading 0.

6. What do you understand by perfect indistinguishability. Suppose you have designed an private key encryption encryption scheme which is perfectly secret. Does this imply that your scheme is perfectly indistinguishable. Justify your answer with complete proof. Consider a private key encryption scheme which is perfectly indistinguishable, and the receiver don't have the decryption key. Will he be able to generate the original message from received cipher-text or not? justify your answer.

7. (a) What do you understand by the *negligible function*? Give it's definition. Give two examples of negligible function and justify.

   (b) Which of the following is/are negligible function(s)? Justify.

      i. $\frac{1}{2^n}$
      ii. $\frac{1}{(\log n)!}$
      iii. $\frac{1}{(\log \log n)!}$
      iv. $\frac{1}{10^{10}}$
      v. $n^{\frac{1}{n}}$
      vi. $\frac{1}{2}$
      vii. $\frac{n}{2^n}$
      viii. $\frac{1}{n}$

   (c) Let f, g be *negligible functions.* Decide whether:

      i. $H(n) = f(n) + g(n)$
      ii. $H(n) = f(n) \times g(n)$
      iii. $H(n) = f(n)/g(n)$

      are necessarily *negligible functions* (for arbitrary f, g) or not. If it is, prove it. If not, give a counterexample.

8. (a) Give a formal definition of the Pseudo random generator. What Do you understand by expansion factor of a Pseudo random generator?

   (b) Say G is a pseudo random generator taking n-bit inputs and producing 2n-bit outputs. Which of the following are necessarily true? (The symbol '|' is used here for string concatenation.) Justify your answer.

      i. G(r) is computationally indistinguishable from a uniform, 2n-bit string if r is a uniform n-bit string.

      ii. G(0 | r) is computationally indistinguishable from a uniform, 2n-bit string if r is a uniform (n-1)-bit string.

      iii. r | G(r) is computationally indistinguishable from a uniform, 3n-bit string if r is a uniform n-bit string.

      iv. G(r) | G(r+1) is computationally indistinguishable from a uniform, 4n-bit string if r is a uniform n-bit string.

      v. G is one-way function.

9. What do you understand by one-way function? Let f, g be length preserving one-way function (so, e.g., $|f(x)| = |x|$). For each of the following functions h, decide whether it is necessarily a one-way function (for arbitrary f, g) or not. If it is, prove it. If not, show a counterexample.

   (a) $h(x) \overset{def}{=} f(x) \oplus g(x)$.

   (b) $h(x) \overset{def}{=} f(f(x))$.

   (c) $h(x_1 \parallel x_2) \overset{def}{=} f(x_1) \parallel g(x_2)$, ($\parallel$ means concatenation)

   (d) $h(x_1, x_2) = (f(x_1), x_2)$ where $|x_1| = |x_2|$.

10. Consider the below Private key encryption scheme: Let G be a pseudorandom generator with expansion factor l. The private key encryption is defined as follow:
Gen: on input 1^n, choose k= $\{0,1\}^n$ uniformaly at random and output it as a key.
Enc: input k and m $\epsilon\{0,1\}^j$, output cipher-text as $c := G(k) \oplus m$.
Dec: input k and c $\epsilon\{0,1\}^j$, output cipher-text as $m := G(k) \oplus c$.
Where j = l(n).
Prove or refute:

   (a) The private key encryption scheme defined above has the indistinguishable encryptions in the presence of an passive adversary (eavesdropper).

   (b) The private key encryption scheme define above has indistinguishable multiple encryptions in the presence of an eavesdropper.

11. Let $G : \{0,1\}^* \rightarrow \{0,1\}^*$ be a function that doubles the length of its input, i.e. $|G(s)| = 2 \cdot |s|$. Show an algorithm A (that does not necessarily run in polynomial time) for which

$$|\mathtt{Pr}[A(G(s)) = 1] - \mathtt{Pr}[A(r) = 1]| \geq 1/2$$

for n large enough. Can we conclude that "perfect PRGs" do not exist, why ?

12. Define the following function G taking n-bit inputs and producing (n+1)-bit outputs: G(x)=x‖0, where ‖ denotes concatenation. Which of the following attackers shows that this G is not a pseudorandom function? (Note: Only one of the below is true.)

    (a) On input an (n+1)-bit string y, output 0 if the first bit of y is 0.

    (b) On input an (n+1)-bit string y, output 1 if the first bit of y is 0.

    (c) On input an (n+1)-bit string y, output 0 if the last bit of y is 0.

    (d) On input an (n+1)-bit string y, output 0 if the first bit of y is equal to the last bit of y.

13. Let F be a pseudorandom function with 128-bit key and 256-bit block length. Which are the following functions G are pseudorandom generators?

    (a) G(x)=$F_x(0...0)\|F_x(0...0)$, where x is a 128-bit input.

    (b) G(x)=$F_{0...0}(x)F_{1...1}(x)$, where x is a 256-bit input

    (c) G(x)=$F_x(0...0)$, where x is a 128-bit input.

    (d) G(x)=$F_x(0...0)\|F_x(1...1)$, where x is a 128-bit input.

14. Given an efficiently-computable function $G : \{0,1\}^* \to \{0,1\}^*$ with $|G(x)| = l(|x|)$ consider the following experiment defined for an algorithm A and parameter n:

    (a) Choose random $s \in \{0,1\}^n$ and set $y_0 = G(s)$. Choose random $y_1 = \{0,1\}^{l(n)}$.

    (b) Choose a random bit $b \in \{0,1\}$.

    (c) Give y_b to A, who outputs a bit b'.

    say G is an *indistinguishable* PRG if for all probabilistic, polynomial-time algorithms A, there exists a negligible function ϵ such that

    $$\mathtt{Pr}[b' = b] \leq \frac{1}{2} + \epsilon(n)$$

    in the experiment above.

    Prove that this definition is equivalent to the definition of a pseudorandom generator.

15. state true or false with brief explanation:

    (a) Any private-key encryption scheme that is CPA-secure must also be computationally indistinguishable.

    (b) Any private-key encryption scheme that is CCA-secure must also be perfectly secret.

    (c) Any private-key encryption scheme that is CCA-secure must also be CPA-secure.

16. Let F be a block cipher with 128-bit block length. Consider the following encryption scheme for 256-bit messages: to encrypt message $M = m_1 \| m_2$ using key k (where $|m_1| = |m_2| = 128$), choose random 128-bit r and compute the ciphertext r $\|F_k$(r) $\oplus$ $m_1 \|F_k(m_1) \oplus m_2$. Which of the following strategies would lead to a valid chosen-plaintext attack? (Note: only one is true.)

    (a) Choose random r and let m be arbitrary but not equal to r. Output messages M_0=$r\|m$ and $M_1 = m\|m$. Output 0 if the second block of the challenge cipher text is all-0s.

    (b) There is no attack; this scheme is randomized, so it is CPA-secure.

    (c) Let m_1 and m_2 be arbitrary but distinct. Using the encryption, obtain an encryption $r\|c1\|c2$ of $m2\|m2$. Output messages $M_0 = m_1\|m_1$ and $M_1 = m_1\|m_2$. Output 0 if the third block of the challenge ciphertext is c_2.

    (d) Let m_1 and m_2 be arbitrary but distinct. Using the encryption , obtain an encryption $r\|c_1\|c_2$ of $m_1\|m_2$. Output messages $M_0 = m_1\|m_2$ and $M_1 = m_2 m_1$. Output 0 if the third block of the challenge Ciphertext is c_2.

17. What do you understand by these terms: one way permutation, Pseudorandom generator, one way function, pseudorandom function, invertible pseudorandom generator. Give complete details (and if possible present an example illustrating the methods you describe) of how to use an X to design a Y where:

    (a) X = One-way permutation, Y = Pseudorandom generator.

    (b) X = Pseudorandom generator, Y = One-way function.

    (c) X = Pseudorandom generator, Y = Pseudorandom function.

    (d) X = Pseudorandom function, Y = Invertible pseudorandom function.

18. Let F be a block cipher with n-bit block length. Consider the following encryption scheme: to encrypt a message viewed as a sequence of n-bit blocks $m_1, m_2, , m_t$ using a key k, choose a random n-bit value r and then output the ciphertext r,$F_k(r + 1 + m_1), F_k(r + 2 + m_2), , F_k(r + t + m_t)$, where addition is done modulo 2n. Which of the following attackers demonstrates that this scheme is not computationally indistinguishable: (Note: Only one is true.)

    (a) Let m be an arbitrary n-bit block, and output M_0=m,m and M_1=m,m1. Given challenge ciphertext r,c_1, c_2, output 1 if and only if $c_1 = c_2$.

    (b) Let m be an arbitrary n-bit block, and output M_0=m and M_1=m,m. Given a challenge ciphertext, output 0 if the challenge ciphertext contains 2 blocks, and output 1 otherwise.

    (c) Choose random n-bit blocks m and m', and output M_0=m,m and M_1=m,m. Given challenge ciphertext r,c_1, c_2, output 1 if and only if $c_1 = c_2$.

    (d) Choose random n-bit blocks m_1, m_2, m_3, m_4, and output $M_0 = m_1, m_2$ and $M_1 = m_3, m_4$. Given challenge ciphertext r, c_1, c_2, output 0 if r= 0....0, and output 1 otherwise.

19. (a) Why probabilistic encryption scheme is required ? When the private key encryption scheme has indistinguishable encryption under chosen- plaintext attack? Explain in Detail. If a private key encryption scheme has indistinguishable encryption under chosen- plaintext attack, does this imply that this private key encryption scheme has indistinguishable multiple encryption under chosen- plaintext attack?

    (b) What do you understad by Pseudo random function? Give formal definition. Assume the pseudorandom function exists, create a CPA-Secure fixed length private key encryption schemes using the Pseudorandom function. Prove how it is CPA-secure. Make proper assumptions if needed.

20. (a) What do you understand by Message integrity? What do you understand by Message Authentication code? Explain the construction of Fixed length MAC and Variable length MAC in Brief.

    (b) Prove that the basic CBC-MAC is not secure when we consider the messages of different length. can you perform modification in Basic CBC-MAC for variable length messages So that it will be secure? Explain in detail.

    (c) Let F be a block cipher with n-bit block length. Consider the message authentication code for 2n-bit messages defined by $Mac_k(m_1, m_2) = F_k(m_1 m_2)$. Which of the following gives a valid attack on this scheme?

        i. Obtain tag t on message m,m, and then output the tag 0...0 on the message 0...0,m.

        ii. Obtain tag t on message $m_1, m_2 (with m_1 != m_2)$, and then output the tag t on the message m_2, m_1.

        iii. Obtain tag t on message m,0...0 (with ! = 0...0), and then output the tag t on the message 0...0,0...0.

        iv. Obtain tag t on message m,0...,0, and then output the tag $t \oplus$ (1...1) on the message m,1...1.

21. (a) What do you understand by Hash function? What properties a Hash function should possess from Cryptographic point of view? Explain the Birthday Attack in brief. Explain these terms: *Encrypt and Authenticate*, *Authenticate then Encrypt*, *Encrypt then Authenticate.*

    (b) Assume we want to use a hash function with output length as small as possible, subject to being collision resistant against a birthday attack running in time $2^1 92$. Which hash function would be the best choice from these functions? SHA-3 with 384-bit output, SHA-1, MD5, SHA-2 with output truncated to 192 bits.

    (c) Let H,H' be collision-resistant hash functions. Which of the following functions H"is NOT necessarily collision-resistant?

        i. $H''(x) = H(x)H'(x)$, where $\|$ denotes concatenation.

        ii. $H''(x) = H(x) \oplus H'(x)$.

        iii. $H''(x) = H(H'(x))$.

        iv. $H''(x) = H(x)\|0...0$, where $\|$ denotes concatenation.