CS402 Cryptography: Worksheet

1. The letter "N" was found to be the most frequent letter in a large ciphertext produced using a Caeser cipher over the alphabet $A = 0, B = 1, ..., Z = 25$. Decipher the following portion of the ciphertext:
$$CQRBRBJFNJTLRYQNA \quad .$$

2. (a) Describe the Vigenère cipher.

   (b) A Vigenère cipher over the 27-letter alphabet $A = 0, B = 1, ..., Z = 25, \_ = 26$ was used to produce the ciphertext

   $EKWWEKKWWEKKDQDIKLOFDTKKSRKCA\_PWKZWQPXJWVK$
   $SWNDDSIKDXE\_\_GKE\_WHKDWRQLVMRQPK\_DMWPHSR$
   $OKQROYPDRMQKQRLLOHDBKNPSRNPHKLLDH\_DNZBCHKX$
   $IBIKDPFWPKXWS\_WMRHLYKWDLYLRQLIJWDDSMPHKMMR$
   $LL\_XPDQZZREKGP\_PPLIS \quad .$

   The corresponding plaintext begins

   $A\_LONG\_LONG\_TIME\_AGO\_ \quad .$

   Determine the first 17 words of plaintext.

3. What basic property of the Vigenère cipher makes it significantly more secure than the Caeser cipher for sending very short messages (such as a bank PIN or credit card details)?

4. What basic property of the Enigma cipher makes it significantly more secure than the Vigenère cipher against a ciphertext only attack.

5. The ciphertext
$$TMY$$
was produced by applying the affine enciphering function $f_E\colon \mathbb{Z}_{26} \to Z_{26}, x \mapsto 9x + 16$ to single letter message units over the alphabet $A = 0, B = 1, ..., Z = 25$. Determine the plaintext.

6. The ciphertext
$$ESDCWNMH$$
was produced by applying a Hill cipher to 2-letter message units over the alphabet $A = 0, B = 1, ..., Z = 25$, with enciphering finction
$$f_E\colon \begin{pmatrix} x \\ y \end{pmatrix} \mapsto A \begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{where } A = \begin{pmatrix} 2 & 3 \\ 7 & 5 \end{pmatrix}.$$
Calculate A^{-1} mod 26 and hence determine the first FOUR letters of plaintext.

7. Carefully explain each of the following italicized terms:

   (a) *Kerckhoff's principle*

   (b) *two factor authentication*

   (c) *symmetric* cipher

   (d) *public key* cipher

   (e) *block* cipher

   (f) *stream* cipher

   (g) *key space*

   (h) *ciphertext only attack*

(i) *known plaintext attack*

(j) *frequency analysis attack*

(k) *computationally secure* cipher

(l) *perfectly secure* cipher

(m) *forward secrecy* of a cipher

(n) *public key signature with message recovery*

8. List the five main attributes/operations of an object oriented implementation of a cryptosystem.

9. Determine, with proof, the size of the enciphering key space for an affine cipher $f_E\colon (\mathbb{Z}_p)^d \to (\mathbb{Z}_p)^d, v \mapsto AV + B$ over an alphabet of p letters with p a prime.

   When working over an alphabet with $p = 29$ letters, what is the smallest value of d for which the cipher could be considered computationally secure against a ciphertext only attack? Justify your answer.

10. Let $\mathbb{P}$, $\mathbb{C}$, $\mathbb{K}$ denote respectively the plaintext space, ciphertext space and key space of a given cryptosystrem. Assuming that each of these spaces is finite, that the cryptosystem is perfectly secure and that every $c \in \mathbb{C}$ has non-zero probability of occuring, prove the inequalities
$$|\mathbb{K}| \geq |\mathbb{C}| \geq |\mathbb{P}|.$$

11. Describe one perfectly secure (though possibly impractical) cryptosystem. By carefully stating and using a theorem of Shannon, or otherwise, explain why the cryptosystem is perfectly secure.

12. (a) Explain what is meant by an L-bit *linear feedback shift register*, and explain how it is represented by its connection polynomial.

    (b) A 4-bit linear feedback shift register with connection polynomial $C(C) = 1 + X + X^3$ is used to produce a pseudo-random sequence of binary digits, staring $s_0 = 0$, $s_1 = 0$, $s_2 = 1$, $s_3 = 1$. By listing the next few terms of the pseudo-rndom sequence, determine its period.

13. The enciphering key for a binary stream cipher is produced using a 3-bit linear feedback shift register. It is known that the cipher converts the plaintext string

    0010001101010111

    into the ciphertext string

    1000010000011001 .

    (a) Determine the first sixteen binary digits in the enciphering key.

    (b) Determine the connection polynomial of the linear feedback shift register.

14. Why is the non-linear function $f(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_3$ a poor choice if used to combine three binary linear feedback shift registers in order to get a "non-linear" pseudo-random sequence.

15. Carefully describe the A5/1 stream cipher which is used to encrypt the on-air traffic in the GSM mobile phone networks in Europe and the US.

16. Diffie and Hellman published a paper in 1976 on *public key cryptography*. The idea had in fact been invented five years earlier, independently and under the name *non-secret encryption*, by James Ellis who worked for the British govenrment's communications headquarters GCHQ – an organization not always eager to publicize its work. What is public key cryptography?

17. The problem of constructing a public key encryption system was given to a new recruit to GCHQ called Clifford Cocks in 1973. Cocks had studied mathematics as an undergraduate at Cambridge and as a postgraduate at Oxford. Within a day at GCHQ Cocks had invented what is essentially the RSA algorithm, a full four years before Rivest, Shamir and Adleman published their public key cryptosystem. Rivest was a mathematics graduate from Yale, Shamir a mathematics graduate from Tel Aviv, and Adleman a mathematics graduate from Berkeley. Describe the RSA cryptosystem.

18. Explain how the RSA cipher can be used as a public key signature with message recovery.

19. Explain how the RSA cipher could be used in two factor authentication (such as the HSBC secure keypad).

20. Alice joins an RSA public-key cryptosystem with public key

$$(n, e) = (713, 7)$$

over a 26-letter alphabet $A = 0, ..., Z = 25$. A plaintext message corresponds to the integer 37. Find the integer corresponding to the enciphered message.

21. Let $N = pq$ with p and q distinct primes, let e be an integer with $gcd(\ e,\ (p-1)(q-1)\ ) = 1$, and let $d \equiv e^{-1} \bmod (p-1)(q-1)$.

    (a) Let m be any integer such that $gcd(m, N) = 1$. Assuming Euler's Totient Theorem, prove that
    $$(m^e)^d \equiv 1 \bmod N.$$

    (b) Indicate the relevance of the above result to RSA cryptography.

    (c) State and prove Euler's Totient Theorem.

22. There exist practical methods for testing the primality of a large integer and yet there is no known practical method for factoring a large integer as a product of primes. How is this situation possible?

23. What is a *pseudo-prime* to the base b? Is 91 a pseudo-prime to the base 2? Is 91 a pseudo-prime to the bae 3?

24. What is a *Carmichael* number?

25. Suppose that m is **not** a pseudo-prime to some base $b \in \mathbb{Z}_N^*$. Prove then that m is not a pseudo-prime to at least half of the possible bases in $\mathbb{Z}_N^*$.

26. Describe Fermat's primality test. Suppose that m is a composite number which is not a Carmichael number. Estimate the probability that m passes the test k times.

27. Carefully describe the *discrete logarithm problem* (DLP) and the *Diffie-Hellman problem* (DHP) .

28. In 1974 an employee at GCHQ and Cambridge mathematics graduate, Malcolm Williamson, invented the concept of *Diffie-Hellman key exchange.* Describe this concept.

29. Alice and Bob choose the abelian group $\mathbb{Z}_{71}^*$ and generator $g = 7$ to perform Diffie-Hellman key exchange. Alice secretly chooses $a = 6$ and Bob secretly chooses $b = 13$. What is the numerical value of their shared key?

30. What is a *man in the middle attack* on the basic Diffie-Hellman key exchange, and how is it resisted using RSA cryptography?

31. Let E denote the set of points on the elliptic curve over $\mathbb{Q}$ defined by $y^2 = x^3 + ax + b$. Describe (without giving precise formulae) the operations of addition and subtraction which give E the structure of an abelian group.

32. Briefy describe Lenstra's elliptic curve factorization algorithm.

33. Use Pollard's rho method, with the given $f(x)$ and x_0, to factorize the following integers n. In each case compare x_k with x_j for which $j = 2^h - 1$ (where $2^h \leq k < 2^{h+1}$).

    (a) $n = 91,\ f(x) = x^2 - 1,\ x_0 = 2$.

    (b) $n = 8051,\ f(x) = x^2 + 1,\ x_0 = 1$.

    (c) $n = 7031,\ f(x) = x^2 - 1,\ x_0 = 5$.