# COMP8760

RSA Practice Worksheet

Sanjay Bhattacherjee

**Note.** Please consult the handout note on RSA provided earlier to solve this worksheet.

**"Naive" RSA.** Fill in the blanks.

1. **Setup.**

   (a) The setup algorithm takes as input the security parameter $\nu$ and generates two _____-bit prime numbers $p$ and $q$.

   (b) Let $N =$ _____. Since $N$ is composite, $\mathbb{Z}_N = \{$_____, _____, _____, $\ldots$, _____$\}$ is a ring with respect to _____ and _____ modulo $N$.

   (c) We also have $(\mathbb{Z}_N)^*$ as the set of all non-zero elements of _____that are mutually prime to _____. It forms a group with respect to multiplication modulo _____.

   (d) The number of elements in this group is given by

   $$|\_\_\_\_\_| = \phi(\_\_\_\_\_)$$

   where $\phi(N)$ is Euler's totient function.

   (e) In particular, we have $\phi(\_\_\_\_\_) = \phi(p)\phi(q) = (\_\_\_\_\_)(\_\_\_\_\_) = \_\_\_\_\_ - (\_\_\_\_\_ + \_\_\_\_\_) + 1$.

   (f) Note that there are _____ multiples of $q$ between 0 and $N-1$ namely those in the set

   $$\{\_\_\_\_\_ \cdot q, \_\_\_\_\_ \cdot q, \ldots, \_\_\_\_\_ \cdot q\}.$$

   (g) Similarly, there are _____ multiples of $p$ between 0 and $N-1$ namely those in the set

   $$\{\_\_\_\_\_ \cdot p, \_\_\_\_\_ \cdot p, \ldots, \_\_\_\_\_ \cdot p\}.$$

   (h) Since 0 is a multiple of both $p$ and $q$, hence the total number of integers outside of $(\mathbb{Z}_N)^*$ is (_____ + _____ $-1$).

   (i) Let $M = \phi(N)$. We will invoke the notation $\phi(N)$ when we need to keep in mind that $\phi(N) =$ (_____ $-1$)(_____ $-1$). Otherwise, we will treat this number only as a composite integer $M$.

   (j) For large values of $p$ and $q$, the fraction $\frac{\phi(N)}{N}$ is close to _____. In other words, the set $(\mathbb{Z}_N)^*$ is almost as large a set as $\mathbb{Z}_N$.

   (k) We recollect that the extended Euclidean algorithm takes as input two integers $a$ and $b$ and finds their *greatest common divisor* $d$ in a form

   $$\_\_\_\_\_ = x \cdot \_\_\_\_\_ + y \cdot b.$$

   The coefficients $x$ and $y$ are also returned by the algorithm. We note here that the extended Euclidean algorithm is a very efficient algorithm.

   (l) The next step is to find an integer $e$ such that $\gcd(e, M) =$ _____. In other words, we will have to find an element $e \in$ _____. We may generate a random integer $x \leq M$ and check if _____ $= 1$. We note here that $|(\mathbb{Z}_M)^*| = \phi(M)$ has to be large to easily find such an $e$.

   (m) Once we have found such an $e$, the output of the extended Euclidean algorithm will be in the form

   $$1 = x \cdot e + y \cdot \_\_\_\_\_.$$

   (n) Applying the modular operation $\pmod M$ on both sides of this equation, we get

   $$x \cdot e = \_\_\_\_\_ \pmod{\_\_\_\_\_}.$$

(o) This element $x$ output by the algorithm is the inverse of _____ in the multiplicative group _____. We denote this inverse element by $d$. So,

$$d \cdot \_\_\_\_\_ = 1 \pmod{\_\_\_\_\_}. \tag{1}$$

(p) In summary, the output of the setup phase is:

public key $\mathfrak{pk} \leftarrow (\_\_\_\_\_, \_\_\_\_\_)$

secret key $\mathfrak{sk} \leftarrow (N, d)$

(q) Note that the primes $p$ and $q$ may only be kept hereafter for improving the efficiency of exponentiation as has been explained in Section 6.3.2 in the book [**?** ]. The value of $\phi(N)$ may not be stored at all.

2. **Encryption.**
The user's message is mapped to an element $m \in \mathbb{Z}_N$. The encryption algorithm takes as input the message $m$ and the public key $\mathfrak{pk} = (N, e)$, and finds the ciphertext $c$ as follows.

$$c \leftarrow \_\_\_\_\_ \pmod{\_\_\_\_\_}.$$

3. **Decryption.**
The decryption algorithm takes as input the ciphertext $c$ and the secret key $\mathfrak{sk} = (N, d)$, and finds the message $m'$ as follows.

$$m' \leftarrow \_\_\_\_\_ \pmod{\_\_\_\_\_}.$$

We say that the system works correctly if the message $m$ that was encrypted is indeed the one that has been found as the output of decryption. In other words, the system works correctly if $m = m'$.

4. **Correctness.**

(a) To show that the scheme is correct, we have to show that

$$\begin{aligned} m = m' &= c^d \pmod{N} \\ &= (\_\_\_\_\_)^d \pmod{N} \\ &= (\_\_\_\_\_)^{e \cdot d} \pmod{N} \end{aligned} \tag{2}$$

(b) At this point, we recollect from Equation 1 that $e \cdot d = \_\_\_\_\_ \pmod{\_\_\_\_\_}$. So, we may write $e \cdot d = \_\_\_\_\_ + s \cdot \_\_\_\_\_$ for some integer $s$.

(c) Substituting this value of $e \cdot d$ in Equation 2, the new statement to be proved is

$$\begin{aligned} & m = m^{\_\_\_\_\_ + s \cdot \_\_\_\_\_} \pmod{\_\_\_\_\_} \\ \implies & m = m \cdot m^{s \cdot \phi(\_\_\_\_\_)} \pmod{\_\_\_\_\_} \\ \implies & 0 = m \cdot \left( m^{s \cdot \phi(\_\_\_\_\_)} - 1 \right) \pmod{\_\_\_\_\_}. \end{aligned} \tag{3}$$

To prove Equation 3, we will show that

$m \cdot \left( m^{s \cdot \phi(N)} - 1 \right)$ *is a multiple of* $N$.

(d) We recollect at this point that $m \in \mathbb{Z}_N$. So,

- either $m \in (\mathbb{Z}_N)^*$,
- or $m \notin (\mathbb{Z}_N)^*$.

In case $m \notin (\mathbb{Z}_N)^*$, then either $m$ is one of the _____ multiples of $p$, or $m$ is one of the _____ multiples of $q$. When $m$ is a multiple of $p$, it is mutually prime with $q$ and vice versa. Only when $m =$ _____, both $p$ and $q$ divide $m$.

(e) In summary, there are four possible scenarios as listed below.

| | divisibility by $p$ | | divisibility by $q$ |
|---|---|---|---|
| Case 1: | $\gcd(m, p) =$ _____ | and | $\gcd(m, q) =$ _____ |
| Case 2: | $m =$ _____ $\pmod{p}$ | and | $\gcd(m, q) =$ _____ |
| Case 3: | $\gcd(m, p) =$ _____, | and | $m =$ _____ $\pmod{q}$ |
| Case 4: | $m =$ _____ $\pmod{p}$, | and | $m =$ _____ $\pmod{q}$ |

We look at each of these cases in the following.

i. **Case 1:** This is the most common case as mentioned earlier. In this case, since $\gcd(m, p) =$ _____ and $\gcd(m, q) =$ _____, hence $m \in (\mathbb{Z}_N)^*$. By Euler's theorem, we have

$$\begin{aligned} m^{\phi(\_\_\_\_\_)} &= 1 \pmod{\_\_\_\_\_} \\ \implies \left(m^{\phi(\_\_\_\_\_)}\right)^s &= 1 \pmod{\_\_\_\_\_} \\ \implies m^{s\cdot\phi(\_\_\_\_\_)} &= 1 \pmod{\_\_\_\_\_}. \end{aligned} \tag{4}$$

Hence, $N$ divides $m \cdot \left(m^{s\cdot\phi(\_\_\_\_\_)} - 1\right)$ and hence Equation 3 is proved to be correct.

ii. **Case 2:** Since $\gcd(m, q) = 1$, we use Fermat's little theorem to note that

$$\begin{aligned} m^{(\_\_\_\_\_-1)} &= 1 \pmod{\_\_\_\_\_} \\ \implies \left(m^{(\_\_\_\_\_-1)}\right)^{(p-1)} &= 1 \pmod{\_\_\_\_\_} \\ \implies m^{\phi(\_\_\_\_\_)} &= 1 \pmod{\_\_\_\_\_} \\ \implies \left(m^{\phi(\_\_\_\_\_)}\right)^s &= 1 \pmod{\_\_\_\_\_} \\ \implies m^{s\cdot\phi(\_\_\_\_\_)} &= 1 \pmod{\_\_\_\_\_}. \end{aligned} \tag{5}$$

From Equation 5 we get that _____ divides $\left(m^{s\cdot\phi(\_\_\_\_\_)} - 1\right)$. We already know that in this case, $p$ divides $m$. Combining these two facts, we get that $N$ divides

$$\underbrace{m}_{\text{divisible by } \_\_\_\_\_} \cdot \underbrace{\left(m^{s\cdot\phi(\_\_\_\_\_)} - 1\right)}_{\text{divisible by } \_\_\_\_\_}$$

and hence Equation 3 is proved to be correct.

iii. **Case 3:** The arguments in this case are the same as Case 2, by interchanging the prime $q$ with $p$.

iv. **Case 4:** For $m = 0$, Equation 3 is trivially correct. However, note that $c = m^e \pmod{N} =$ _____ which is unchanged from $m$.

5. **Note on the choice of $m$.**

(a) In practice, for any choice of $N = p \cdot q$, there will be (_____ + _____ − 1) elements $m \in \mathbb{Z}_N$ for which $\gcd(m, N) > 1$. These are precisely the elements $m \notin (\mathbb{Z}/N\mathbb{Z})^*$ and hence $\gcd(m, N) \neq 1$.

(b) If an attacker of the system (with only the knowledge of the public key $\mathfrak{pk} =$ (_____, _____)) is able to find such a message $m$, then they can compute $\gcd(m, N)$ to get one of the two primes $p$ or $q$. This will provide the factorisation for $N$ and as a result, the cryptosystem will be broken.

(c) Now, if all messages $m \in \mathbb{Z}_N$ are equally likely to occur, then the probability that an attacker gets hold of such a message $m$ for which $\gcd(m, N) > 1$ is negligibly small. The precise probability is given by

$$\frac{\_\_\_\_\_}{N}.$$

(d) We note that $p$ and $q$ are $\nu/2$-bit primes. So, $p + q - 1$ would be of around that size as well while $N$ will be approximately of size $\nu$ bits. The probability of finding an element $m \notin (\mathbb{Z}_N)^*$ is around

$$\_\_\_\_\_.$$

(e) Hence, Case _____, Case _____and Case _____of the correctness proof are extremely unlikely to occur. However unlikely, if an innocent sender of messages actually takes up one such $m \notin (\mathbb{Z}_N)^*$ and tries to encrypt it, the encryption and decryption mechanisms as designed, will work correctly. That is because of the cases above where we have shown that the correctness property still holds in these two cases.