NATIONAL UNIVERSITY OF SINGAPORE

CS4236 - CRYPTOGRAPHY THEORY AND PRACTICE
(Feedback)
(Semester 1: AY2022/2023)

Time allowed: 2 hours

INSTRUCTIONS TO STUDENTS

This assessment paper contains **FIVE (5)** sections totalling **FORTY** (**40**) marks, and comprises **TEN (10)** printed pages including this one.

This is an **OPEN BOOK** assessment, and you are to answer **ALL** questions. You may cite any result in the lecture notes or tutorials. Answer **ALL** questions within the space provided in this booklet (write on the backs of pages if you need more room).

Please write your Student Number below. (Do not write your name).

STUDENT NO: _____

This portion is for examiners use only.

Question		**Marks**	**Remark**
General topics, short answers	Q1 (8)		
HASH and MAC	Q2 (8)		
Symmetric encryption	Q3 (8)		
Asymmetric encryption, signatures	Q4 (8)		
Secrets and MPC	Q5 (8)		
Total:	Q1-5 (40)		

Q1 (Short Answer Questions) (8 marks)

In the following short questions, each answer is worth either 1 (ONE) or 2 (TWO) marks.

1.1 Calculate the bias for $x \oplus y \oplus z$ (with x, y, z independent) when $\varepsilon(x) = 0.1$, $\varepsilon(y) = 0.2$, and $\varepsilon(z) = 0.3$. Show your working. (2 marks)

In this case you should indicate you are using the piling up lemma and have the correct answer - $\varepsilon(x \oplus y \oplus z) = 2^2 \times 0.1 \times 0.2 \times 0.3 = 0.024$

1.2 Explain what type of oracle is a padding oracle, and what sort of attack is one that uses a padding oracle? (2 marks)

A padding oracle is a (partial) decryption oracle and the attack is a CCA.

1.3 Calculate the entropy in bits/symbol, of a source emitting the 5 symbols E, X, A,M and S, with the probability of E being 0.5, and the other four having equal (0.125) probabilities. Show your working. (2 marks)

The entropy is 2 bits/symbol, and I expect to see you use the equation as given in the course.

1.4 Give a (variable length) bit encoding for each of the characters E,X,A,M and S derived from the source in Q1.3, which would have an average bit length/symbol the same as that which you just calculated in Q1.3. (1 mark)

Something like E=0, X=100, A=101, M=110, S=111 would work. (1)

1.5 In (say) the symmetric EAV-security game (Defn 3.8), the adversary is PPT. Explain why it must be PPT - what would happen with an unbounded adversary? (1 mark)

If it was unbounded, then the adversary could just try all 2^n keys.

Q2 (HASH and MAC) (8 marks)

2.1 Prove that MD5 is not collision resistant. (2 marks)

> In class I mentioned MD5 in a few different contexts, but always mentioned that there are known collisions. So you can start with the argument that "There exists at least one MD5 collision". I was expecting some argument that there exist collisions, so if an adversary has one, it can always win the "collision resistant" game by producing that example.

2.2 Assume you had a hash function (like MD5), which is known to NOT be collision resistant, but may be 2nd preimage resistant. Outline (describe) a construction that could turn a 2nd preimage resistant hash function into a collision resistant hash function. (3 marks)

> Understanding of the question is enough to get one point, but for the rest I am looking for some idea - perhaps to add a random value to every hash, and add it to the hash for canonical verification.

2.3 In the above construction, a sender might do the $\text{Mac}_k(m)$ to generate the tag, and a remote receiver might do $\text{Vrfy}_k(m,t)$ to check that the message is undamaged. However, a man-in-the-middle could just modify the message and regenerate the tag. Outline how you could ensure that this would not happen. (3 marks)

> Understanding of the question is enough to get one point, but for the rest I am looking for something like: the keys must be private to the sender and receiver. Some of you may develop more subtle/complex systems - perhaps even using SIG systems...

Q3 (Symmetric encryption) (8 marks)

3.1 The above SPN has a localized permutation. Briefly explain why it is not as useful as an SPN. (2 marks)

In your answer you should give an example of a specific attack on the SPN, and estimate the (brute force) attack time, if each trial took 0.001 *second.*

Understanding of the basic issue gets a point, and a specific attack might be to do 16 trials, putting in the (same) values 0000 to 1111 in each input S-box. This will provide a complete input-output mapping in 0.016 second.

3.2 Briefly explain why in CCA-secure symmetric encryption systems, the encryptions cannot be deterministic. Describe what strategy is commonly used to make symmetric encryption systems non-deterministic. (3 marks)

If the encryption was deterministic, then an adversary will be able to win a choice game. The use of IVs and modes is normal to make them non-deterministic.

3.3 Shown above is the linear approximation table for the toy SPN example used in class. To the right are two stages of linear cryptanalysis, with particular inputs and outputs highlighted. Calculate the bias of the top and bottom dots (at the k_2 and k_4 levels). I have shown the relevant intermediate interconnections. (3 marks)

In your answer you should show the bias for each of the relevant S-boxes (T_1, T_2), showing how you used the approximation table, and also calculate the final bias (top-to-bottom).

The bias for T_1 is $\frac{1}{4}$ (as $N_L(c,4) = 12$). The bias for T_2 is $\frac{1}{8}$ (as $N_L(4,a) = 10$).Final bias is $\frac{1}{16}$

Q4 (Asymmetric encryption, signatures) (8 marks)

4.1 Explain why encrypting with a private key using (textbook) RSA is not existentially unforgeable. In particular - what attack can be done? (3 marks)

In your answer you should give an example of a specific attack that defeats the Sig-forge *game, briefly explaining what property of textbook RSA leads to the attack.*

Recognition that textbook RSA is homomorphic wrt multiplication gets one mark. For the rest I expect an attack based on creating a new (multiplied) message.

4.2 In the Diffie-Hellman key agreement protocol, the two participants (Alice and Bob say) broadcast the values $g^a \bmod p$ and $g^b \bmod p$, and can then compute $g^{ab} \bmod p$. Unfortunately, this requires more messages if we had three participants - Alice, Bob and Charles. Explain the messages needed if three participants do Diffie-Hellman key agreement. (3 marks)

In your answer you should outline each message, what it is, who it is from/to...

This was not a question where you can use bilinear maps. I expect you to outline the messages using normal DHKA. There are two rounds. In the first round A sends g^a to B, B sends g^b to C, C sends g^c to A. B, C and A can now calculate g^{ab}, g^{bc} and g^{ca} respectively. Finally A sends g^{ca} to B which calculates g^{cab}, B sends g^{ab} to C which calculates g^{abc}, and finally C sends g^{bc} to A which calculates g^{bca}. These values are all the same: the key. There may be other solutions. (3)

4.3 In many asymmetric encryption schemes, messages to be sent are encoded as a bit string. This is divided into fixed size blocks which are treated as (very big) integers (in $\mathbb{Z}$), and then various critical mathematical operations are applied to these integers. ECC and some other encryption schemes are based on cyclic groups that have elements not in $\mathbb{Z}$ (ECC for example uses points in a plane). As a result messages are encoded as points on a plane, and the critical mathematical operations are applied to those points. However, it need not be that way. Briefly describe a technique that allows the messages to still be treated as (very big) integers (in $\mathbb{Z}$), but with the critical mathematical operations still being done on the plane. (2 marks)

Understanding of question gets a point. A technique would involve a hash function from points on the plane to points $\mathscr{H} : \mathscr{G} \rightarrow \mathbb{Z}_n$, as we saw in the last lecture.

Q5 (Secrets and MPC) (8 marks)

5.1 In the secret-sharing construction (above left) given in Topic13, the claim is made that this is SS-secure (above right). Prove it. (5 marks)

In your answer you may refer specifically to the example given, where all secrets are in the range $0 \ldots 10^6 - 1$*, and (if needed) 3 participants.*

Proof should follow style of proof in class (direct, suppose not, justification). I expect the idea that knowing other shares does not change the probability for winning - it is still just $\frac{1}{10^6}$.

5.2 An evaluator in a Yao garbled circuit MPC scheme is able to find the correct output label specific to a particular set of input labels. Explain exactly how the evaluator does this for a particular circuit, and why the evaluator may still have no idea of the value attached to that label. (3 marks)

In your answer you could use the AND gate example above.

The evaluator uses the two input labels as keys to decrypt the supplied AND value, returning k_k^b. Because the evaluator may not know the label/value table $\mathcal{T}_{w_k}$, it does not know if this represents a 0 or a 1.

=== END OF PAPER ===