IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2015-2016

MEng Honours Degree in Mathematics and Computer Science Part IV
MEng Honours Degrees in Computing Part IV
MSc in Advanced Computing
MSc in Computing Science (Specialist)
MRes in High Performance Embedded and Distributed Systems
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the Associateship of the City and Guilds of London Institute

PAPER C408H

PRIVACY ENHANCING TECHNIQUES

Wednesday 16 December 2015, 14:00
Duration: 70 minutes

Answer TWO questions

Paper contains 3 questions
Calculators not required

1a. What problem does the Millionaires' problem solve? How can the solution be adapted for equality?

b Consider the following 1-from-n oblivious transfer protocol in an honest-but-curious (semi-honest) model.

1. Alice generates n random public-private key pairs

   $(pub_1, priv_1), \ldots (pub_n, priv_n)$

   Alice sends the public keys $pub_1, \ldots, pub_n$ to Bob.

2. Bob generates n random symmetric keys $k_1, \ldots, k_n$ and computes

   $G_b = E_{pub_b}(k_b)$ and $G_z = k_z$ for all $z \in \{1..n\}$ and $z \neq b$

   Bob sends G_1 to G_n to Alice

3. Alice computes

   $H_z = D_{priv_z}(G_z)$,

   $C_z = E_{H_z}(M_z)$ for all $z \in \{1..n\}$

   Alice sends C_1 to C_n to Bob

4. Bob computes $M_b = D_{k_b}(C_b)$

For this protocol:

i) Explain why Bob's output equals M_b.

ii) Explain why Alice learns nothing about b. What assumptions do you have to make about the two cryptosystems for this to be true?

iii) Explain why Bob learns nothing about M_z for $z \neq b$. What assumptions do you have to make about the two cryptosystems for this to be true?

iv) If Alice were dishonest, is there anything she could do to learn b? If so, describe how. If not, explain why not.

v) If Bob were dishonest, is there anything he could do to learn messages other than M_b? If so describe how. If not, explain why not.

c What is a Garbled Table? Illustrate your answer for XOR.

The three parts carry, respectively, 15%, 60% and 25% of the marks.

2a Give a formal definition of an ε-differentially private mechanism.

b What does parameter ε mean? What values can it take and what effect do they have on the privacy mechanism?

c Researcher Ray is conducting a survey and will be using a differential-privacy mechanism for the results. Ray makes the following statements about privacy to potential participants. Using an example explain why each statement is invalid under differential privacy:

i) An adversary won't be able tell whether or not you participated in the survey and won't be able to learn anything about you from his results.

ii) An adversary won't be able to reliably determine whether you took the survey.

d Consider a survey with the following 3 questions:

Have you read a Harry Potter book (Yes/No)?

How many books in the Harry Potter series do you own (0 to 7)?

What is your gender (Male, Female)?

State and briefly explain the global sensitivity for each of the following differentially private analyses on the collected Harry Potter dataset.

i) How many people have read a Harry Potter book?

ii) How many males *and* how many females are in the dataset?

iii) An analysis consisting of both of the previous queries i.e. both (i) and (ii).

iv) An analysis consisting of the following 4 queries. How many males have read a Harry Potter book? How many males have not read a Harry Potter book? How many females have read a Harry Potter book? How many females have not read a Harry Potter book?

v) The total number of Harry Potter books owned?

vi) The average number of books owned?

e Given a dataset that holds whether a person is a graduate or not (using 0 for False, and 1 for True), devise a differentially-private mechanism for a function that returns the percentage of graduates in the dataset. You can assume that the size of the dataset is public knowledge.

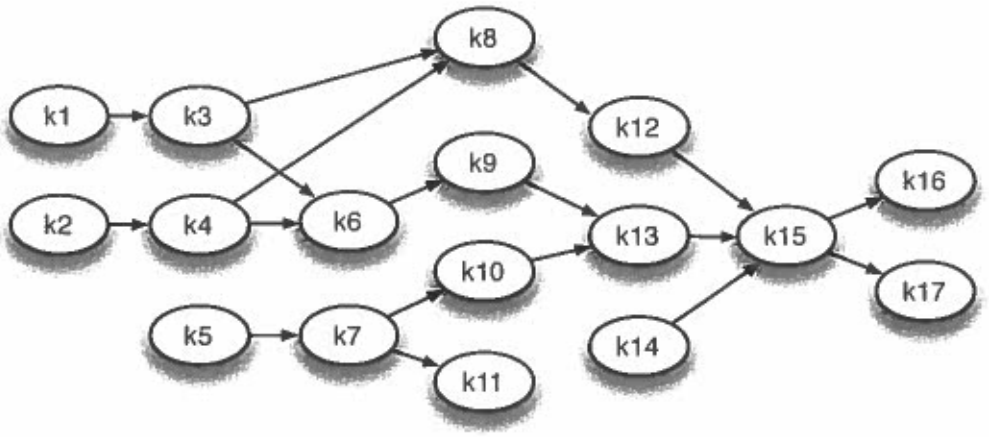
The five parts carry, respectively, 15%, 20%, 20%, 30% and 15% of the marks.

3a Give three examples of when someone might want to use Bitcoins anonymously. What are the risks to anonymity in buying Bitcoins from an Online Exchange?

b Identify 4 capabilities that an expert computer adversary could use to de-anonymise Bitcoin users.

c Most Bitcoin users use browsers to interact with websites to obtain bitcoins, to store them, to send and receive them, and to sell them. Why is this risky and how can you mitigate against it?

d Consider the following Bitcoin user network. Each vertex corresponds to a Bitcoin address and the edges correspond to Bitcoin transactions. For this user network how many users are there? Explain your answer.



e What types of analysis that can be performed on the user network to deduce information about Bitcoin users? How can a network analyst map IP addresses to Bitcoin addresses?

The five parts carry, respectively, 30%, 20%, 20%, 10%, 20% of the marks.